Structural and Functional Characterisation of Human Cdc7 Kinase

Samual David Dick

University College London

and

The Francis Crick Institute

PhD Supervisor: Dr. Peter Cherepanov

A thesis submitted for the degree of

Doctor of Philosophy

University College London

January 2017

Declaration

I, Samual David Dick, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Cell division cycle 7-related Ser/Thr (Cdc7) kinase is conserved and essential across eukaryotes. When bound to its activator, dumbbell-forming factor 4 (Dbf4) it phosphorylates a number of target proteins involved in various aspects of the cell cycle, including replication initiation, meiosis, the intra S-phase checkpoint, the DNA damage response and mitotic exit. Cdc7 is overexpressed in a number of cancers and expression correlates with patient prognosis. Selective inhibition of Cdc7 leads to cell death through an aberrant S-phase in transformed cells, while healthy fibroblasts are able to survive such treatments. Therefore, Cdc7 is an attractive target for cancer therapeutics, and high-resolution structural information could be very informative for the development of small molecule inhibitors.

Herein, I present crystal structures of a fully active Cdc7-Dbf4 heterodimeric construct bound to a potent Cdc7 inhibitor, a non-hydrolysable ATP analogue and an Mcm2-derived substrate peptide. These structures, refined to a high resolution, reveal a previously unseen Zn-binding domain that supports a fully open conformation of the active site. The structures also reveal features required for substrate binding and phosphorylation. *In vitro* assays have been employed to validate the functional significance of these features, and an *in cellula* system developed to investigate their importance in the cell. The new structural and functional information gained from this study will inform the design of small molecule inhibitors of Cdc7 while the cell-based system opens up new ways of addressing the functions of the unique kinase insert sequences of Cdc7.

Acknowledgement

I would firstly like to thank my supervisor, Dr. Peter Cherepanov for the opportunity to spend the last four years working on such a technically diverse and interesting project and for his support and guidance throughout. I would also like to thank my thesis committee Dr. John Diffley and Dr. Alessandro Costa for their input and useful discussion of my work. A special thanks also goes to John for allowing me to use his illustrator files for production of the origin licensing and origin firing figures in this thesis. A number of people from the many wonderful core facilities I have had at my disposal throughout my PhD have contributed to this thesis, including Nicola O'Reilly and the peptide synthesis lab, without whom much of this project would not have been possible. I would also like to thank the cell services team and the FACS facility for supplying me with cells and advising me on a number of experiments that were integral to the outcome of this project. I would also like to thank Vesela Encheva of the proteomics facility who helped in the design and analysis of a number of mass spectrometry experiments. David Bacon and Corella also provided helpful feedback and advice on the content and construction of this thesis. I would also like to give a huge thank you to all of the lab aids and support staff at Clare Hall and the Crick for supporting me in my work and allowing me to purely focus on science.

I have been very fortunate in my time as a PhD student to find myself surrounded by a close-knit group of incredibly friendly and talented scientists. I would particularly like to thank all members past and present of the Chromatin Structure and Mobile DNA lab. Special thanks must go to Siobhan for giving me a crash course in all things Cdc7 and making me feel so welcome in my first few months at Clare Hall. I would also like to thank Val for her seemingly endless patience when teaching me the fundamentals of crystallography and navigating Linux systems. Special thanks must also go to Nicola, without whom I'm fairly certain the whole lab would fall apart. Paul and Dan also provided endless entertainment and support, which kept me sane and made my time in the lab all the more enjoyable.

A special mention must also go to my fellow PhD students Paolo, Aga, Melisa, and Kortyna, with whom I survived Chandos Avenue and Clare Hall. I'm pretty sure we can now survive anything! I would also like to thank Diana for being a constant source of positivity at Clare Hall and for the many lifts home she gave me in her frankly terrifying car!

Outside of the lab I have been lucky to meet a number of friends who have grown to become family. With them I have done countless things that I would never have previously thought possible. Their constant support and friendship have made these years some of the best of my life. While there are far too many names to mention, special thanks must go to Joe, Colin, Kealan, Jackie and Matthew. As well as Brook for organising my social life for several years and making my time in London relentlessly fun. I would also like to thank Khalid, not only for his encyclopaedic knowledge of DNA replication, but also for introducing me to the London Frontrunners and the many friends I have made through the group. If I had known joining a running club could have such a positive influence on my life, I would have done it years ago!

Most of all I would like to thank my family for their constant love and support. My sisters Kelly and Holy, for keeping me grounded during my time in London. My father, Graham, for teaching me to always aim high and experience as much of the world as possible and my mother Diane, for her love, support and courage in the face of adversity, which has always been a constant source of inspiration. All of my achievements, including this PhD, stem from the great start in life and total support they have given me. I cannot say thank you enough!

Table of Contents

Table of figures

List of tables

Abbreviations

ACN	acetonitrile
ADP	adenosine 5’-diphosphate
AMP-PCP	β,γ-methyleneadenosine 5′-triphosphate
AMP-PNP	Adenosine 5′-(β,γ-imido)triphosphate
APC/C	anaphase promoting complex/ cyclosome
APS	ammonium persulphate
ARS	autonomously replicating sequence
ASK	activator of S-phase kinase
ATP	adenosine 5’-triphosphate
ATPγS	adenosine 5’-O-(3-thio)triphosphate
ATR	ataxia telangiectasia and Rad3 related
BCA	bicinchoninic acid
BME	*β*-mercaptoethanol
bp	base pairs
BRCA1	breast cancer 1
BRCT	BRCA1 C-terminus
BSA	bovine serum albumin
Cas9	CRISP associated endolnuclease
CCD	charge coupled device
Cdc	cell division cycle
CDK	cyclin-dependent kinase
Cdt	Cdc10 dependent transcript
Chk1	checkpoint kinase 1
Clb	cyclin B
CMG	Cdc45-MCM-GINS complex
CRISPR	clustered regularly interspaced short palindromic repeats
Csm3	chromosome segregation in meiosis 3
Dfp1	Dbf in *pombe* 1
DMEM	Dulbecco’ modified Eagle’s medium
Cryo-EM	cryo-electron microscopy
Erk	extracellular signal–regulated kinase

ES cells	embryonic stem cells
Dbf	dumbbell former
DCM	dichloromethane
DDK	Dbf4-dependent kinase
DIC	diisopropylcarbodiimide
DMF	dimethylformamide
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
Dpb	DNA polymerase B possible subunit
Drf1	Dbf4-related factor 1 (aka Dbf4B)
DSB	double strand break
dsDNA	double stranded DNA
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
eIF	eukaryotic initiation factor
EM	electron microscopy
FACS	fluorescence associated cell sorting
FBS	foetal bovine serum
FL	full-length
GF	gel filtration
GFP	green fluorescent protein
Gwl	greatwall kinase
GINS	go, ichii, ni, san complex
Him1	Hsk interacting molecule 1
His_6	hexahistidine tag
HP1	heterochromatin protein 1
HRP	horseradish peroxidase
HRV14	human rhinovirus 14
HU	hydroxyurea
IPTG	isopropyl beta-D-1-thiogalactopyranoside
IRK	insulin receptor kinase
Kb	kilobase
KI	kinase insert
LC-MS	liquid chromatography-mass spectrometry

MAPK	mitogen activated protein kinase
Mcm	minichromosome maintenance
MEF	mouse embryonic fibroblast
MEN	mitotic exit network
Mer2	meiotic recombination protein 2
mRNA	messenger ribonucleic acid
MW	molecular weight
NCMR	non-conserved middle region
NES	nuclear export sequence
NLS	nuclear localisation signal
NP40	Nonidet P40
NRS	nuclear retention sequence
NTP	nucleoside triphosphate
OCM	ORC-Cdc6-MCM
OCCM	ORC/Cdc6/Cdt1/MCM
ORC	origin recognition complex
PAGE	polyacrylamide gel electrophoresis
PBD	polo box domain
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PEG	polyethylene glycol
PES	polyethersulfone
PFU	plaque-forming unit
PK	protein kinase
PMSF	phenylmethylsulfonyl fluoride
Pol	polymerase
PP1	protein phosphatase 1
PP2A	protein phosphatase 2A
pre-IC	pre-initiation complex
pre-LC	pre-loading complex
pre-RC	pre-replicative complex
Psf	partner of Sld5
PVDF	polyvinylidene fluoride

RecQL	RecQ protein-like
Rfm1	repression factor of MSEs protein 1
Rif1	Rap-interacting factor 1
RIPA	radioimmunoprecipitation assay
RPA	replication protein A
RPC	replication progression complex
rpm	revolutions per minute
Scc2	sister chromatid cohesion protein 2
SDS	sodium dodecyl sulphate
Sld	synthetically lethal with *dpbb11-1*
ssDNA	single stranded DNA
Sum1	suppressor of *mar 1-1*
TEMED	tetramethylethylenediamine
TFA	trifluoroacetic acid
TGS	Tris-glycine-SDS buffer
TICRR	TopBP1-interacting checkpoint and replication regulator
TIS	triisopropylsilane
TLS	translesion synthesis
Tof	topoisomerase I associated factor
TopBP	topoisomerase II binding protein
WT	wild-type

Chapter 1. Introduction

1.1 Overview of eukaryotic protein kinases

The protein kinase complement of the human genome, also known as the kinome, represents a large and functionally diverse collection of proteins. The group consists of 538 different genes which corresponds to approximately 1.7% of all expressed genes in the genome (Manning et al., 2002). This large group can be divided into a number of related groups based upon their sequence similarity with 478 of these forming a single family with highly conserved sequences required for the catalytic activity of the kinase. This group can then be further divided into 7 groups (Figure 1.1). A further 40 kinases belong to an atypical group which are predicted to have similar folding to a typical kinase but are more divergent in their sequence (Manning et al., 2002). Cdc7, the focus of this study, belongs to the CMGC family of kinases. This group is comprised of 9 highly conserved kinase families, most notably the mitogen activated protein kinases (MAPKs) and the cyclin dependent kinases (CDKs), which play integral roles in the cells response to external stimuli and progression through the cell cycle respectively (Schaeffer and Weber, 1999, Malumbres, 2014). Even within the CMGC group, the structure of the kinases vary widely with CDKs being characterised by the requirement of a cyclin for activation of the kinase while MAPKs such as ERK2 are monomeric and can therefore activate without the binding of an accessory protein (Canagarajah et al., 1997). CMGC kinases are often proline directed, requiring a proline residue in the P+1 site of the target sequence, although this feature is not strictly conserved across the entire group of kinases (Echalier et al., 2010). This vast number of structurally diverse kinases allows them to be involved in nearly all aspects of cellular biology including but not limited to cell cycle progression, transcription, translation and metabolism despite all catalysing the same reaction.

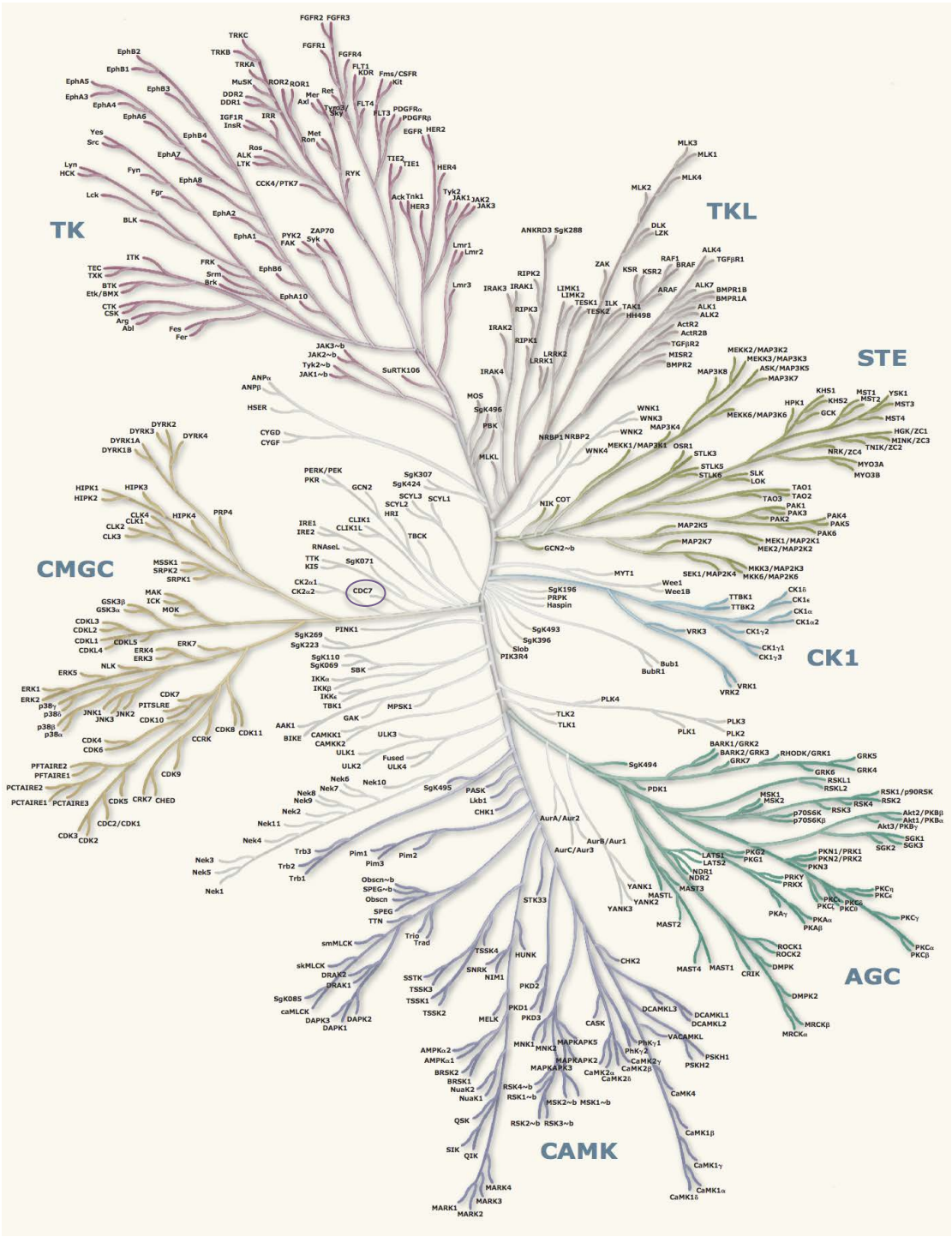


Figure 1-1 A phylogenetic tree of the protein kinase complement of the human genome

(Manning et al, 2002) The position of Cdc7 is circled.

1.1.1 Kinase function

The core catalytic function of a kinase is the transfer of a phosphoryl group (the γ-phosphate of ATP) to a phosphor-acceptor. In the case of protein kinases this acceptor is a serine, threonine or tyrosine residue in a target protein. Typically kinases are either serine/threonine specific or tyrosine specific in terms of their preferred target sites. A small group of dual specificity kinases have also been identified which can phosphorylate any of the possible target residues (Lindberg et al., 1992). Phosphorylation represents the most abundant and arguably most important post-translational modifications in the cell (Sefton and Shenolikar, 2001). Addition of phosphoryl groups to target proteins can lead to a number of outcomes including; the activation or inhibition of enzymatic activities, the localisation of proteins to specific cellular compartments, the creation of binding sites for protein-protein interactions as well as the stabilisation of proteins or marking them for subsequent degradation. Consequently kinases are involved in almost all aspects of cellular function (Cohen, 2002).

In a typical catalytic cycle, the kinase binds to the nucleotide (ATP) and any required divalent metal ions before binding the target substrate. The interaction between a kinase and its substrate is often fairly transient to facilitate the immediate release of the target protein once the phosphoryl transfer has successfully taken place. The nucleotide is then released and it is the release of the nucleotide which is often the rate limiting step of the reaction (Adams, 2001). However, further studies have suggested that the concentration of divalent metals utilised for catalysis can also be rate limiting. An example of this is CDK2, for which two magnesium ions are required for efficient phosphoryl transfer. However, binding of the second magnesium stabilises the binding of ADP after the reaction takes place, leading to a slower turnover rate for the reaction (Jacobsen et al., 2012). It is clear that the various factors that can affect kinase activity have been fine tuned throughout evolution to strike a balance between efficiency of phosphoryl transfer and the turnover of the kinase. In the final stage of the reaction, ADP is released from the active site, allowing the kinase to become competent for a new cycle of catalysis.

1.1.2 Conserved structural features of kinases

Despite their diverse structures and targets, kinases all possess a number of core structural features required to facilitate catalysis (Figure 1.2). The structural basis for the importance of such features was first shown by the crystal structure of PKA in complex with a peptide inhibitor and ATP, closely followed two years later by the first crystal structure of CDK2 (Knighton et al., 1991a, Knighton et al., 1991b, De Bondt et al., 1993). Countless structures of other kinases have since confirmed the essential nature of a number of these features, and structures of kinases in various active and inactive conformations have revealed the numerous conformational changes and post-translational modifications required for successful kinase activation (Endicott et al., 2012). Kinases typically appear most similar to each other in their ‘on’ state and a greater degree of conformational plasticity is observed prior to their activation in which the active site does not have to be correctly formed for catalysis (Huse and Kuriyan, 2002).

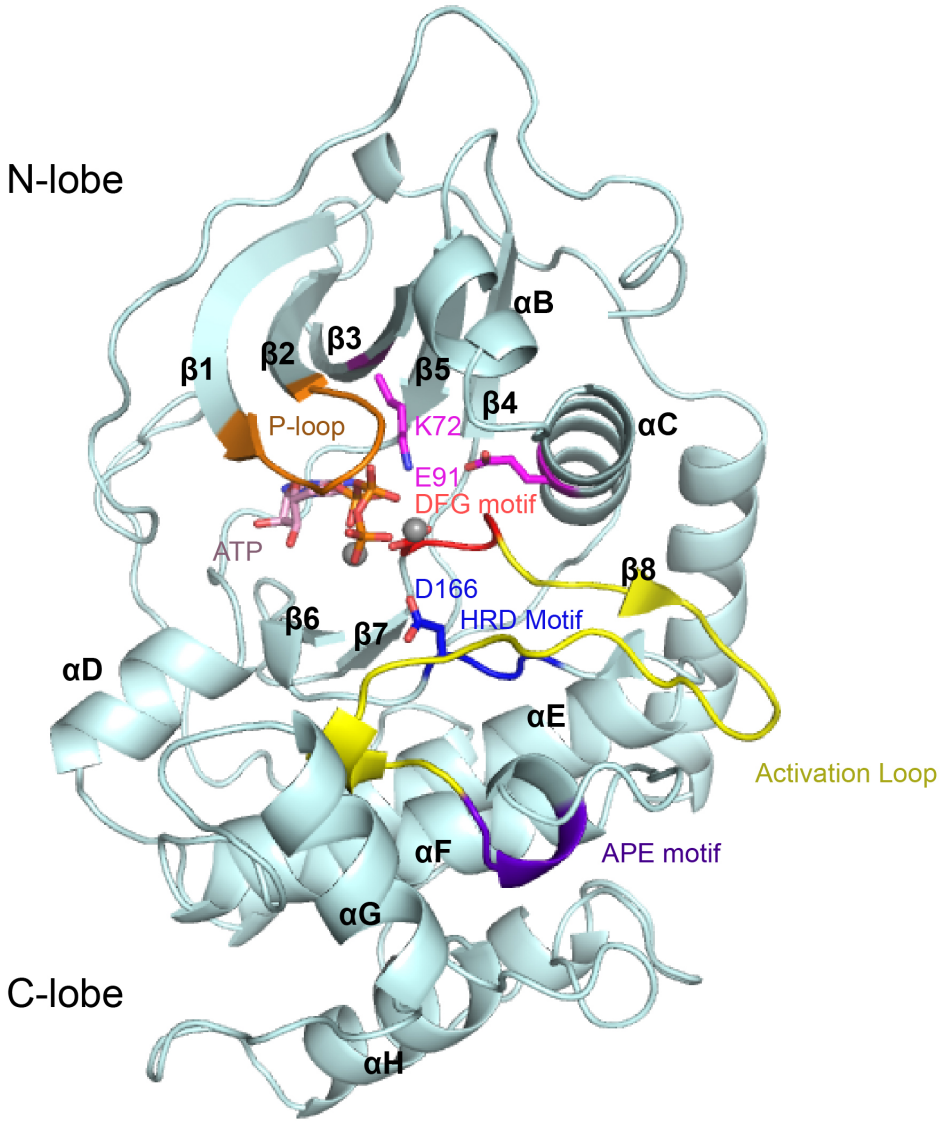


Figure 1-2 Main structural features of a protein kinase

Schematic of protein kinase A is shown as a carton in pale cyan with key structural features highlighted (PDB: 1ATP). ATP is shown as pink sticks and Mn^{2+} atom are shown as grey spheres. Secondary structural elements are labelled from N to C terminus. The P-loop is shown in orange, the DFG and APE motifs (which define the start and finish of the activation loop) are highlighted in red and purple respectively with the activation loop in yellow. The HRD motif is shown in blue. Key residues involved in catalysis are shown as sticks including the invariant K72 and E91 residues, which form a critical salt bridge and are shown in magenta. D166 represents the catalytic base for catalysis. All non-carbon atoms are shown in standard colouration.

A typical kinase has a bi-lobal architecture with the active site present in a deep cleft between the two lobes. The N-terminal lobe is predominantly comprised of a 5-strand β-sheet and a critical α-helix involved in kinase regulation. The C-lobe

tends to be larger than the N-lobe and is comprised predominantly from α-helices. The two lobes are connected by a short hinge region which, as well as allowing the two lobes to rotate with respect to each other, makes important interactions with the adenosine base of ATP during nucleotide binding. The correct orientation of the lobes is required for activation of the kinase as this allows for the correct formation of the active site at the cleft. This can be achieved through interactions with regulatory proteins such as cyclins or even by N-or C-terminal extensions of the kinase, as seen with AGC kinases in response to phosphorylation events (Jeffrey et al., 1995, Canagarajah et al., 1997, Kannan et al., 2008).

Key features of the N-lobe include the P-loop and the α C-helix. The P-loop is a short loop that connects the β1 and β2 strands of the N-lobe. The region is glycine rich, typically containing three glycine residues, This loop forms the upper region of the ATP binding site and backbone interactions between the P-loop and the α, β and γ phosphates positions the phosphates in the correct orientation for phospho-transfer as well as promoting ADP release and subsequent binding of ATP after catalysis (Kornev and Taylor, 2010, Taylor and Kornev, 2011). The conformation of this loop is indicative of the activation state of the kinase, as alternative conformations prevent the efficient binding of ATP. Phosphorylation of the P-loop can also be used to prevent kinase activity as a means of regulation (Morgan, 1995). The αC-helix is often the only α-helix present in the N-lobe and plays a critical role in kinase activation. A key conformational change during kinase activation is the movement of this helix towards the active site, facilitating the formation of a critical salt bridge between and invariant glutamate residue in the helix and a lysine side chain present in the β3 strand. This lysine is also part of a conserved motif termed the AxK motif. This not only stabilises the correct positioning of a number of structural features required for catalysis, but also allows the Glu and Lys residue to coordinate the α and β phosphates of ATP in the correct orientation for catalysis. Such movement of the αC-helix as a method of kinase activation has been shown in a number of kinase structures (Jeffrey et al., 1995, Yang et al., 2002, Bayliss et al., 2003, Sessa et al., 2005, Taylor and Kornev, 2011, Hughes et al., 2012). In the inactive conformation, movement of the helix out of the active site prevents slat bridge formation and subsequent phospho-transfer.

The C-lobe is typically much larger than the N-lobe and is predominantly made up of α-helices and 4 short β-strands. Loops between these strands have structural significance. The loop between β6 and β7 contains an HRD motif in which the Asp residue is the catalytic base which accepts a proton from the target Ser/Thr during phospho-transfer. This loop is known as the catalytic loop. The loop between β8 and β9 of the kinase is called the activation segment, which spans from the conserved DFG motif to the APE motif. This region can vary quite significantly between kinases in length and sequence. A number of kinases have significant insert sequences in this region, the functions of which are poorly understood and are difficult to elucidate through structural studies (Nolen et al., 2004, Hughes et al., 2012, Ocasio et al., 2016). The activation segment is found between the N and C-lobes of the kinase and despite the extensive differences across kinases, the conformation of this region is essential to its efficient activation. The segment often contributes to the formation of a substrate-binding platform required for substrate recognition. The activation loop is also often the site of phosphorylation events with a number of kinases containing a critical threonine residue that must be phosphorylated for kinase activation. This phosphorylated residue forms part of a charged bond network, which often includes the arginine of the HRD motif and a positively charged residue from the αC-helix. This stabilises the correct active site conformation for phospho-transfer (Adams, 2003). Adjacent to the phosphorylation site is the P+1 loop, which is thought to stabilise the residue in the in P+1 position of the target site to ensure correct orientation of the substrate for phospho-transfer. The DFG and APE motifs that flank the activation segment are also highly conserved, with the DFG motif playing a particularly important role in kinase activity. The aspartate residue of this motif directly coordinates one of the metal ions required for catalysis while the phenylalanine binds in a pocket, which forms part of one of the hydrophobic spines of the kinase. The location of this conserved phenylalanine residue is a reliable indicator of whether the kinase is in an active conformation when the kinase has a ‘DFG-in’ conformation (Treiber and Shah, 2013).

An important result of conformational changes achieved through kinase activation is the formation of the catalytic and regulatory spines of the kinase. These two spines are highly conserved across a large number of kinase structures and were

identified by local spatial pattern alignment. The spines consist of a number of residues that make hydrophobic stacking interactions, which span the N and C lobes of the kinase. Both spines are anchored to the αF-helix. In the case of the regulatory spine, this is achieved through the formation of a salt bridge between a highly conserved aspartate residue and the histidine residue of the HRD motif. A number of conserved hydrophobic residues from both the N and C lobes also contribute to the spine including the phenylalanine of the DFG motif. The 'DFG-in' conformation ascribed to active kinases involves the phenylalnine being correctly positioned to complete the regulatory spine. Inactive kinases will often adopt a 'DFG-out' conformation in which the spine is not correctly formed (Kornev et al., 2008, Treiber and Shah, 2013). The catalytic spine is similarly formed of a number of conserved hydrophobic residues anchored to the αF-helix including the alanine from the previously described AxK motif. The Catalytic spine also includes the adenine ring of the ATP as part of the stacking interactions. The identification of these integral spines has shed some light on how mutations far from the active site are able to influence nucleotide affinity in kinases (Kornev et al., 2008).

1.1.3 Mechanism of phospho-transfer

The process of phospho-transfer can be achieved through two different mechanisms termed associative and dissociative pathways. The most common pathway involves the formation of a dissociative transition state in which the bond between the β and γ-phosphate of ATP is broken prior to the formation of the bond between the phosphoryl group and the target hydroxyl (Adams, 2001). The associative pathway involves the formation of a bond between the γ-phosphate and the hydroxyl group either prior to, or at the same time as hydrolysis of the β-γ phosphate bond.

The critical residues for catalysis can be seen in Figure 1-3 utilising the active site of PKA as an example (Adams, 2001). The dissociative reaction involves nucleophillic attack of the hydroxyl group of the substrate with aspartate of the HRD motif acting as a catalytic base to allow for transfer of the phosphoryl group to the

target hydroxyl. The catalytic base (Asp166) coordinates the proton of the target hydroxyl group orienting the substrate for nucleophillic attack in which its lone pair of electrons is directed through the γ-phosphate to the β-γ bond. This breaks the bond between the γ phosphorous and the bridging oxygen. The negative charge created during this reaction is stabilised by the presence of the divalent metal ions, which are coordinated by the aspartate of the DFG motif (Asp184) and (in the case of PKA) a conserved asparagine residue (Asn 171). The positive charge of a conserved lysine residue (Lys168) also contributes to stabilise the γ-phosphate leaving group during transfer. The proton from the attacking hydroxyl group is simultaneously transferred to the catalytic base (Asp166), which will later be removed to facilitate another round of catalysis.

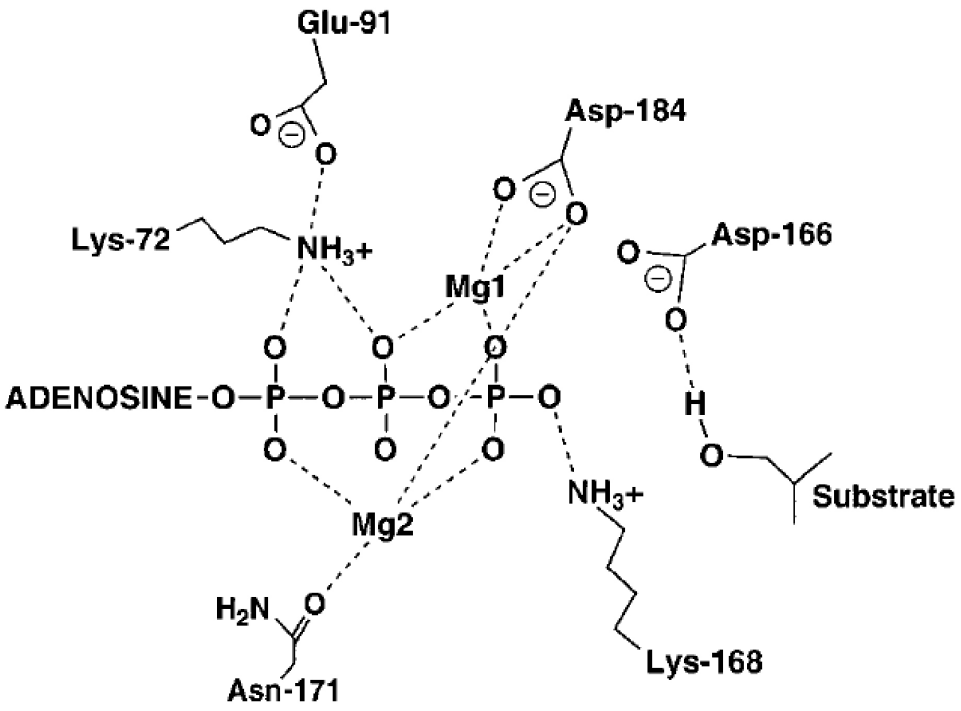


Figure 1-3 Schematic of the active site of a typcial kinase

The shown example is the active site of PKA, taken from Adams, et al. 2001.

1.1.4 Kinase substrate recognition

Kinases vary greatly in the number of targets they are able to phosphorylate in the cell and this number can vary over orders of magnitude as demonstrated by phosphoproteomics analysis in yeast (Ptacek et al., 2005). The phosphorylation of a number of sites in a specific fashion requires kinases to not only be selective but to balance the need for a high level of substrate affinity with a requirement to maintain a fast turnover of phosphorylation reactions. Rapid mixing experiments have shown that the intrinsic affinity of kinases for their substrate is much lower than is suggested by affinities measured by steady state kinetic assays. The increased affinity between the substrate and kinase is due to a feature of the phosphoryl transfer reaction that draws the substrate towards the kinase. This method allows for substantial improvements in affinity while also favouring a fast turnover of sites after the reaction has occurred (Adams, 2001)

Kinases show poor affinity for free amino acids and rely heavily on flanking residues for increased affinity. These local sequence elements are usually only a few (typically no more than 4) residues either side of the target site and often define a consensus sequence for the kinase for which specific residues determine its specificity. Such consensus sequences have been identified for a number of kinases and the length and specificity of such sequences varies between them. Such optimal sequences have been identified through both manual-screening methods in which phosphorylation sites for the kinase are identified and a corresponding library of peptides produced. Substitutions at various points in the sequence are made and the effect of such substitutions on steady state kinetic parameters can be measured. This labour intensive approach can then reveal the importance of each residue in the local sequence (Chan et al., 1982, Pearson and Kemp, 1991). Alternatively, random library approaches have been used in which all possible amino acids, excluding the phosphorylation site (the P site), are altered in the sequence and peptide sequencing and statistical methods can then be used to select optimal peptides (Luo et al., 1995, Songyang et al., 1995, Lam, 1998). The requirements for the consensus of kinases can also be specific to individual groups. An example of this is the CMGC group of kinases, many of which as described previously are proline-directed kinases that favour a proline residue directly

downstream from the target site (Clark-Lewis et al., 1991, Gonzalez et al., 1991, Songyang et al., 1996). In some cases an overlap is seen between kinases in *in vitro* assays highlighting the importance of distal recognition elements in dictating kinase specificity (Songyang et al., 1994, Nishikawa et al., 1997).

Binding determinants away from the active site play an important role in dictating substrate specificity *in vivo* and many kinases also show an affinity orders of magnitude greater for more complete substrates than their corresponding peptides *in vitro* (Hawkins et al., 2000). One such example of a distal recognition element is the recruitment peptide of CDK2; this contains an RXL motif, which binds in a hydrophobic pocket of CDK2 40 Å away from the target peptide. This increases the affinity of the kinase for the peptide (Schulman et al., 1998, Brown et al., 1999b). In many cases substrate-docking sites exist as part of the kinase away from the active site. Such sites function to increase the local concentration of the substrate through an increase in affinity and may serve to activate the kinase allosterically or simply correctly orient the target peptide in the active site (Biondi and Nebreda, 2003). Docking sites have been identified in a diverse range of kinases including 3-phosphoinositide dependent kinase 1 (PDK1), mitogen activated protein kinases (MAPK) and c-Jun N-terminal kinases (JNK) (Kallunki et al., 1994, Biondi et al., 2000, Sharrocks et al., 2000). Beyond such docking sites, specificity of kinases can also be regulated by large-scale protein-protein interactions as well as through the localisation of the kinase in the cell.

Currently very little is known about the substrate specificity of Cdc7. The consensus sequence appears to be any residue with a pre-phosphorylated Ser/Thr or and acidic residue directly downstream from the target site (in the P+1 position) (Montagnoli et al., 2006). As of now the structural basis of this consensus sequence is not known and any further requirements for substrate specificity have yet to be seen. Understanding how Cdc7 binds to its substrates will form a large part of the work in this study.

1.2 DNA replication

An essential function required for the successful survival and growth of any species is the ability to faithfully duplicate its genome, allowing the correct genetic information to be transferred to the next generation of cells. As part of this process, it is important to ensure that the genome is not only correctly duplicated in its entirety with minimal mistakes, but is also only replicated once per cell cycle. Incorrect replication or over-replication can lead to a number of problems including chromosomal instability and rearrangements, which are the driving force for tumorigenesis (Arias and Walter, 2007, Blow and Gillespie, 2008). DNA replication is therefore regulated at a number of levels to ensure correct duplication in eukaryotes, whose genomes are larger and more complex than those of bacteria.

1.2.1 Regulation of DNA replication

In eukaryotes, replication of DNA is strictly coordinated with progression of the cell cycle (Figure 1-4). Genome duplication occurs in S-phase and is largely confined to this period in the cycle by the relative activities of a number of cell cycle kinases and the anaphase promoting complex/cyclosome (APC/C) E3 ubiquitin ligase. Due to the size of the genome, replication has to initiate at multiple origins to allow the full complement of chromosomes to be replicated in a timely fashion. This is in contrast to bacterial replication, which typically initiates from a single origin. DNA replication initiation at eukaryotic origins occurs in a two-step mechanism of origin licensing followed by origin firing (Blow, 1993, Diffley et al., 1994, Remus and Diffley, 2009). Rao and Johnson obtained the first evidence of this two-step mechanism in 1970, by performing experiments involving the fusion of mammalian cells at different points in the cell cycle. Fusion of a cell in G1 with a cell in S-phase led to the initiation of DNA replication in the G1 cell, while fusion of the S-phase cell with a cell in G2 had no such effect. This inferred the presence of a specific factor in S-phase cells that was required for the initiation of DNA replication (Rao and Johnson, 1970). Since this discovery a number of factors have been identified that ensure the sequential action of the origin licensing and origin firing steps. These factors and their roles in regulating DNA replication are discussed in detail in the following sections.

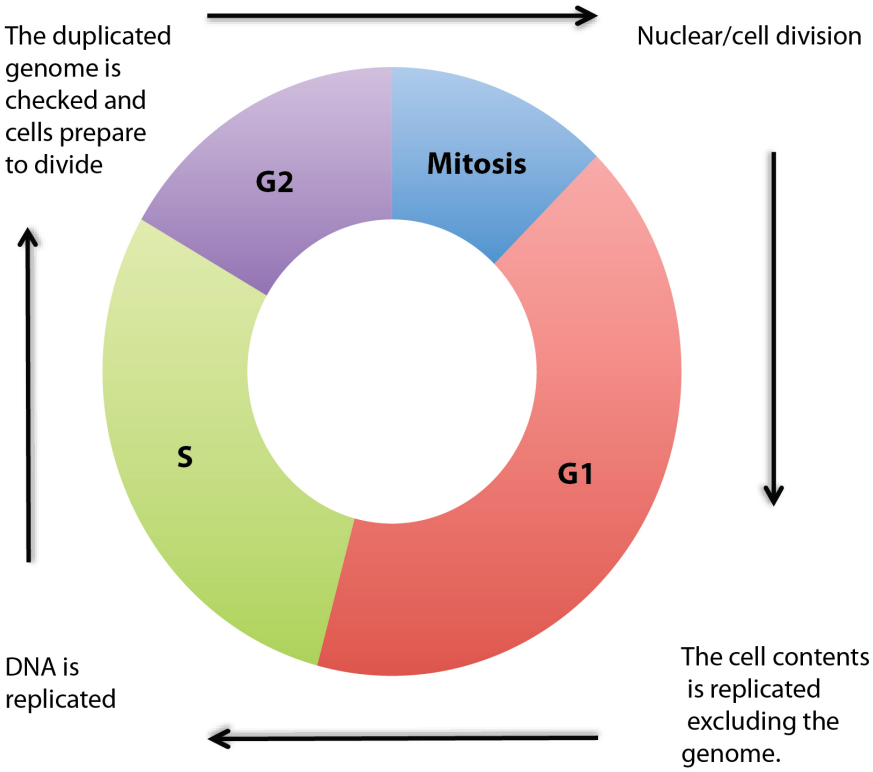


Figure 1-4 Phases of the cell cycle

The cell cycle comprises four distinct phases. In G1, the cell duplicates the organelles and proteins required for the production of two daughter cells. If successful, the cells progress into S-phase, in which the genome is replicated, creating a pair of sister chromatids from each chromosome. In G2, the newly replicated chromatids are checked for errors, and the cell prepares to divide. Lastly, the cell enters mitosis, during which sister chromatids are separated into two nuclei, after which the cytoplasm is split in a process called cytokinesis. This process results in the production of two daughter cells.

1.2.2 Origin licensing

Origin licensing refers to the process that occurs in late mitosis/G1, in which the eukaryotic replicative helicase is loaded onto DNA, ready for subsequent replication. This heterohexameric helicase, known as the minichromosome maintenance 2-7 (Mcm 2-7) complex is responsible for the unwinding of duplex DNA to create the single stranded DNA (ssDNA) templates for loading of the replicative DNA polymerase and DNA synthesis. This process involves a number of factors, which work in sequence to load a pair of Mcm2-7 complexes around double stranded DNA (dsDNA) in an inactive form (Remus et al., 2009, Gambus et al., 2011).

Much of what is known today about origin licensing was determined through a reconstituted system for *in* vitro loading of Mcm2-7 (Evrin et al., 2009, Remus et al., 2009). The first stage in licensing involves the binding of the origin recognition complex (ORC) to a replication origin. The origins are less defined in eukaryotes than in bacteria, where a sequence termed the *oriC,* which contains a specific sequence recognised by the DnaA initiator protein is required for the initiation of DNA replication (Yasuda and Hirota, 1977, Fuller and Kornberg, 1983). Amongst the majority of eukaryotes, replication origins are not defined by specific DNA sequences. In contrast, chromatin modifications, which affect the higher order structure of the genome, appear to be determinants of origin selection in eukaryotic systems. Chromatin can be modified in a number of ways, including the acetylation of histones (Eberharter and Becker, 2002). Acetylation of histone tails leads to a more relaxed and accessible chromatin structure. In *Drosophila* follicle cells hyperacetylation of specific regions of the genome co-localises with newly activated origins of replication, which are made more accessible to the replication machinery to enable developmental transitions (Aggarwal and Calvi, 2004). Chromatin remodelling due to such modifications can also lead to changes in DNA topology, such as the induction of supercoiling. Negative supercoiling has been shown to increase the affinity of the origin recognition complex (ORC) for origins in *Drosophila (Remus et al., 2004).* The budding yeast *S. cerevisiae* is somewhat unusual in that it contains relatively defined origin DNA sequences, termed autonomously replicating sequences (ARSs), which have been shown to act as replication initiation sites *in vivo*. (Brewer and Fangman, 1987). This is one of the many features that made this organism an excellent model for the study of eukaryotic DNA replication. However, histone modifications and nucleosome composition were also shown in *S. cerevisiae* to be important in origin selection. Thus, acetylation of histones H3 and H4 around the origin is required for its activation in S-phase, suggesting the ARS alone is not sufficient for correct origin selection in a physiological setting (Eaton et al., 2010, Unnikrishnan et al., 2010).

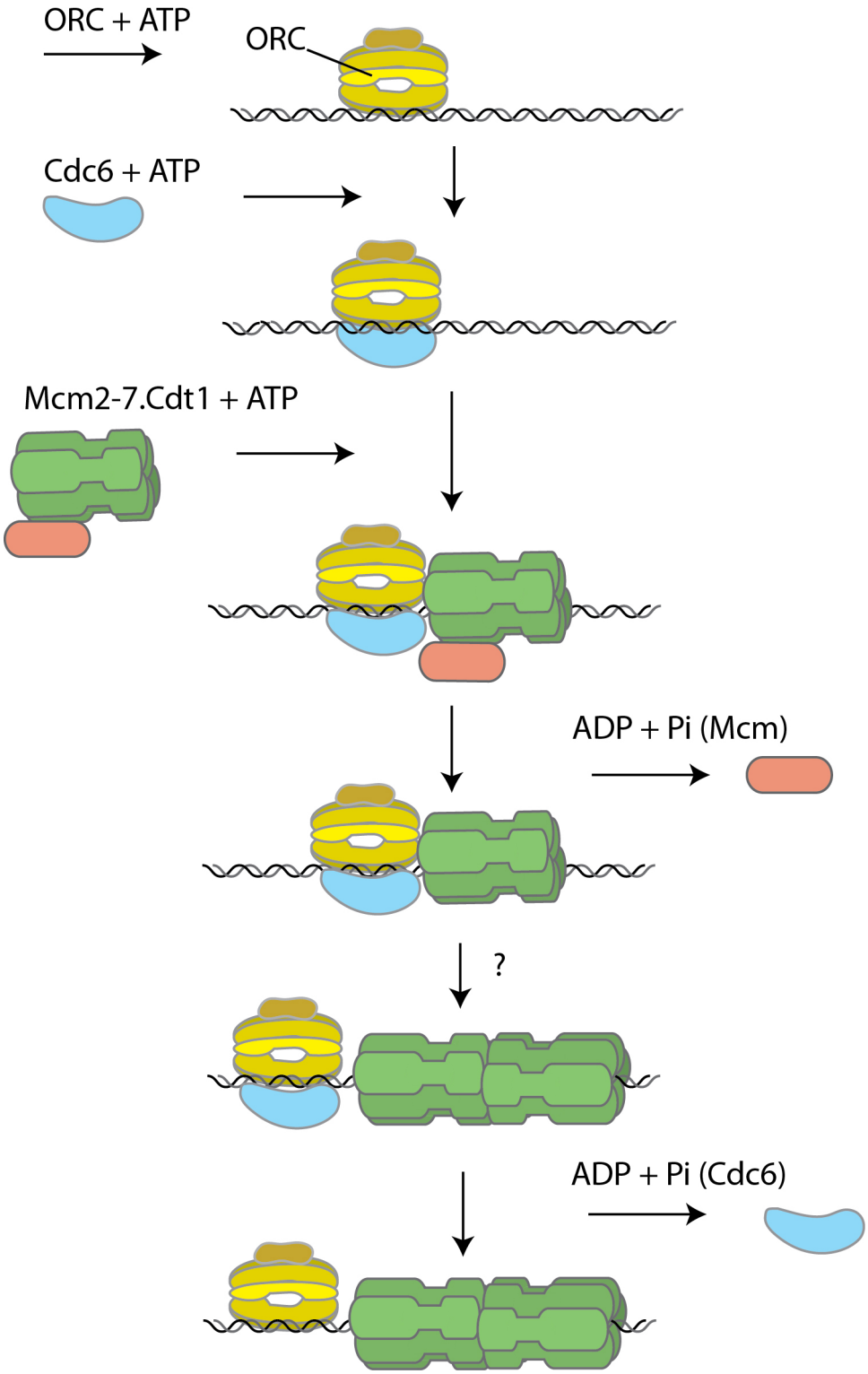


Figure 1-5 A model for the mechanism of origin licensing in *S. cerevisiae*

Proteins referred to in brackets are responsible for ATP hydrolysis at the given stage. For details of the mechanism of licensing see section 1.3.2. (Adapted from Coster et al. 2014).

ORC is composed of six structurally related subunits, Orc1-6, of which two (Orc1 and 5) are able to bind ATP and one (Orc1) is capable of ATP hydrolysis *in vitro.* The ATP binding is essential for ORC binding to origins of replication (Bell and Stillman, 1992, Klemm et al., 1997). While the exact role of ATP hydrolysis by ORC is unknown, recent data indicated that it may not be essential for its Mcm loading activity (Samson et al., 2013, Coster et al., 2014). In licensing, ORC initially binds DNA in an ATP-dependent manner and subsequently recruits cell division cycle 6 (Cdc6) protein, which is also bound to ATP. Engagement of Cdc6 induces a conformational change in ORC, stabilising the ORC-DNA interaction (Speck et al., 2005). The Cdc6-ORC-DNA complex is then competent for the recruitment of MCM, which in most model systems (excluding archaea) is pre-bound to Cdc10-dependent factor 1 (Cdt1) (Tanaka and Diffley, 2002, Ferenbach et al., 2005, You and Masai, 2008). At this point, Mcm2-7 subunits are also bound to ATP, although Mcm2-7 loading does not require ATP hydrolysis (Frigola et al., 2013). This interaction triggers ATP hydrolysis of ORC and Cdc6 and loading of the Mcm2-7 hexamer can occur, forming the pre-replication complex (pre-RC). The process can only occur successfully in the presence of all the pre-RC components and in the absence of CDK. CDK activity, as well as down regulating Cdc6 and promoting exclusion of Mcm2-7 from the nucleus, prevents the loading of the helicase (Nguyen et al., 2001). If these criteria are not met, the ATPase activity of ORC-Cdc6 leads to the release of failed loading intermediates preventing aberrant origin licensing (Frigola et al., 2013).

After successful loading, Cdt1 is the first protein to be released as a result of ATP hydrolysis by the Mcm2-7 complex (Fernandez-Cid et al., 2013, Coster et al., 2014, Kang et al., 2014). Left behind, the complex of ORC, Cdc6 and Mcm2-7 (OCM) is then able to load a second Mcm2-7 complex in a head-to-head orientation. This is thought to occur through a concerted mechanism, as there are no single hexamers observed on DNA when the reaction is carried out in the presence of ATP *in vitro* (Evrin et al., 2009, Remus et al., 2009). However, the mechanism of loading of the second Mcm2-7 remains a mystery, and several models have been proposed (Yardimci and Walter, 2014). Recent single molecule studies suggested that recruitment of the second Mcm2-7 complex occurs through interactions between the two hexamers. Loading of the second hexamer requires distinct Cdc6 and Cdt1

molecules but is directed by the single ORC complex. Interactions between the two hexamers for loading also ensure that loading is bi-directional (head-to-head) (Ticau et al., 2015). After both Mcm2-7 hexamers are loaded, the ATPase activity of Cdc6 is then required for its disengagement, giving way to the further steps in DNA replication initiation (Chang et al., 2015). At the end of origin licensing, a double hexamer of Mcm2-7 is loaded onto dsDNA in a head-to-head conformation, such that the ring formed by N-terminal domains in one interacts with the ring of N-terminal domains of the other, with C-terminal domains pointing outwards (Remus et al., 2009). A recent high-resolution cryo-EM structure of the yeast Mcm2-7 double hexamer from a G1 chromatin-bound fraction revealed fine details of this assembly (Li et al., 2015a). The double hexamer formation was also observed in a metazoan system, suggesting the mechanism of pre-RC assembly is conserved across eukaryotes (Gambus et al., 2011). A model for origin licensing is shown in figure 1-5.

1.2.3 Origin firing

Activation of the loaded replicative helicase occurs at the G1/S-phase boundary and throughout S-phase and requires extensive re-modelling of the Mcm2-7 complex through the recruitment of numerous firing factors. A number of these have been identified in budding yeast including Sld3 (synthetically lethal with *dpbb11-1*), Sld7, Dpb11 (DNA polymerase B possible subunit 11), Sld2 and Mcm10. Furthermore, Cdc45 and the hetero-tetrameric GINS (go-ichii-ni-san) complex, which is composed of Sld5, Psf1 (partner of Sld5), Psf2, and Psf3, are required to interact with the Mcm2-7 complex to form a functional replicative helicase, minimally containing Cdc45, Mcm2-7 and GINS (referred to as the CMG complex) (Gambus et al., 2006, Moyer et al., 2006, Ilves et al., 2010). The molecular transactions leading to origin firing are regulated by cyclin-dependent kinases (CDKs) and dumbbell forming factor 4 (Dbf4)-dependent kinase Cdc7 (Cdc7-Dbf4, often also referred to as DDK) that facilitate or restrict this process (Boos et al., 2012, Tanaka and Araki, 2013). While little is known about the exact mechanisms of Mcm2-7 activation, a number of experiments have elucidated some essential steps required to convert the inactive Mcm2-7 helicase to an active CMG

complex. These steps were largely elucidated through *in vitro* experiments using a cell extract system (Kubota et al., 2003, Pacek and Walter, 2004, Moyer et al., 2006, Heller et al., 2011). More recently, Diffley and colleagues have been able to reconstitute the entire process of origin firing and DNA synthesis *in vitro* using purified components, which required 16 individual replication factors comprised of 42 polypeptides (Yeeles et al., 2015). A schematic of origin firing can be seen in Figure 1-6.

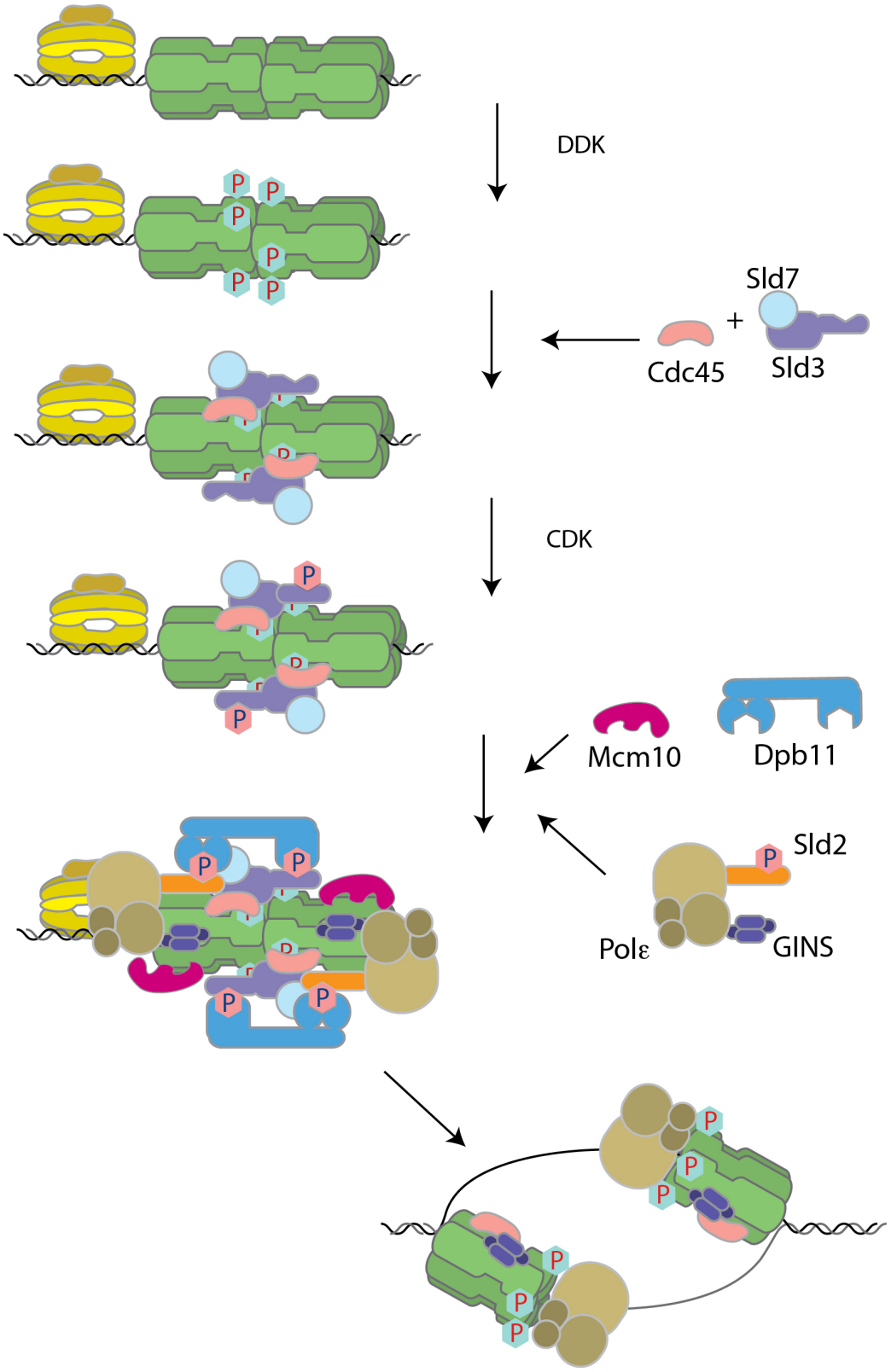


Figure 1-6 A model for the mechanism of origin firing in *S. Cerevisiae*

Phosphorylation events by Cdc7-Dbf4 and CDK are indicated with blue and pink hexagons, respectively. For details of origin firing see section 1.1.4.

The first step in origin firing, the recruitment of Sld3-7 and Cdc45, is dependent on phosphorylation of Mcm2-7 by Cdc7, an S-phase specific Ser/Thr protein kinase. Cdc7 strictly requires an activating subunit and two paralogous Cdc7 activators have been described: Dbf4 and the Dbf4-related factor 1 (Drf1), of which the latter is restricted to metazoans and remains largely unstudied (Yabuuchi et al., 2006, Tanaka et al., 2011, Yeeles et al., 2015). The official name of Drf1 is Dbf4B, however, for clarity, Drf1 will be used throughout this thesis. CDK is not required at this early stage, and it was previously shown that Cdc7-Dbf4 acts upstream of CDK in origin firing (Jares and Blow, 2000, Walter, 2000). However, at least *in vitro,* CDK and Cdc7-Dbf4 are able to act in any order (Yeeles et al., 2015). At the G1/S-phase transition, an increase in CDK activity facilitates the recruitment of Sld3 and Sld2 to the adaptor protein Dpb11. The interaction occurs due to the CDK-dependent phosphorylation of Sld2 and 3 which leads to the recruitment of these proteins to the two BRCA1 C-terminus (BRCT) tandem repeats in Dpb11. Sld3 binds to BCRT I and II while Sld2 is recruited to BRCT III and IV. Simultaneous to this, the pre-loading complex (pre-LC) is also formed, which consists of Dpb11, Sld2, GINS and the eukaryotic leading strand polymerase Polε (Muramatsu et al., 2010). This complex is thought to be relatively unstable. Due to Sld2 and Sld3 being bound to the pre-LC and pre-RC respectively, Dpb11 acts as an adaptor that allows for the CDK-dependent recruitment of GINS to the replisome. Upon full formation of the CMG at the origin, significant re-modelling of Mcm2-7 must occur to allow for DNA unwinding to occur. While recruitment of the various firing factors has been well studied, we still know very little about how the activation occurs. The Mcm2-7 remodelling leads to the formation of two individual CMG complexes that can fire bi-directionally. There is a suggested role for Mcm10 downstream of this formation, which may promote the ATP hydrolysis activity of the helicase and origin unwinding (van Deursen et al., 2012). It is still unknown whether the Mcm2-7 complex encircles ds- or ss-DNA during unwinding, and what triggers the separation of the double hexamers and the initiation of DNA unwinding.

1.2.4 The role of kinases in the regulation of DNA replication

Kinases play an integral role in the regulation of the cell cycle and their expression and activities are tightly coordinated with their functions (Figure 1-7). A number of kinases are involved in regulating the initiation of DNA replication through the phosphorylation of numerous replication factors including the Mcm2-7 subunits.

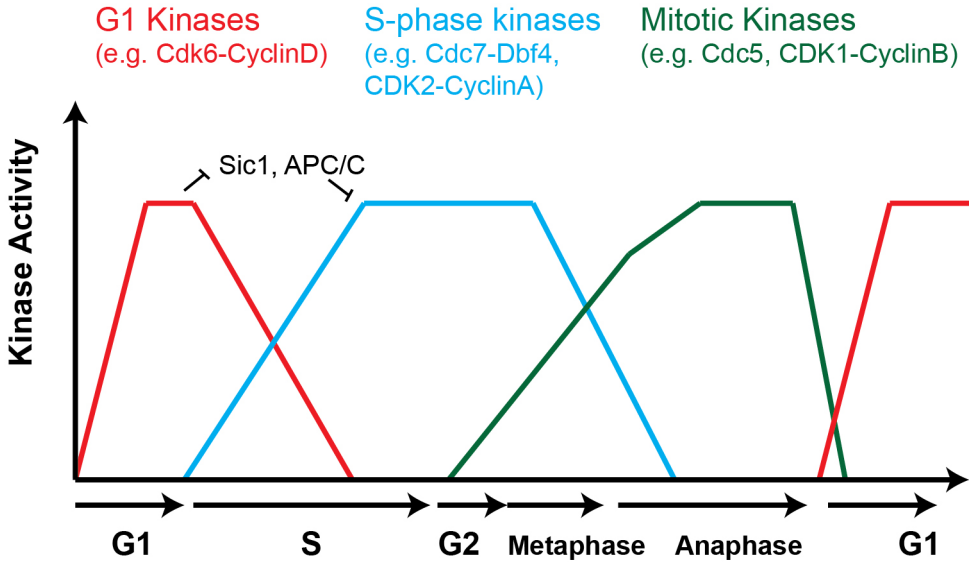


Figure 1-7 Oscillations of kinase activities during the cell cycle

1.2.4.1 Regulation of origin licensing by S-phase CDKs

Restriction of DNA replication to S-phase in the cell cycle can be largely attributed to differences in kinase activity brought about by differential expression of activating subunits of CDKs and Cdc7-Dbf4. In budding yeast, CDK activity as a means of preventing pre-RC formation during S-phase is well studied. S-phase CDKs (Cdc28-Clb5 and Cdc28-Clb6 in yeast) phosphorylate a number of components of the pre-RC to prevent loading, and the cyclin Clb5 can directly interact with the Orc6 subunit (Weinreich et al., 2001, Wilmes et al., 2004). Both forms of yeast S-CDK phosphorylate Orc2 and 6. Phosphorylation of Orc6 specifically leads to a loss of interaction between ORC and Cdt1, preventing pre-RC formation (Nguyen et al., 2001, Chen and Bell, 2011). Phosphorylation of ORC can also lead to the release of helicases prior to successful loading upon ATP hydrolysis by ORC and Cdc6 as part of a proofreading mechanism to prevent

aberrant origin licensing. (Frigola et al., 2013). Moreover, CDKs can also prevent pre-RC formation through phosphorylation of Cdc6, targeting it for degradation by the 26S proteasome. During mitosis, binding of the mitotic cyclin Clb2 to Cdc6 prevents its activation during this stage of the cell cycle (Drury et al., 2000, Mimura et al., 2004). Most of the Mcm2-7 subunits are spatially separated from the origins outside of G1, as CDK-dependent phosphorylation of Mcm3 promotes the nuclear export of the complex (Nguyen et al., 2000, Mimura et al., 2004).

In metazoans, pre-RC formation is additionally regulated by phosphorylation of Cdt1 by CDK, directing this essential Mcm2-7 loading factor for ubiquitination and proteasomal degradation (Li et al., 2003, Nishitani et al., 2006). However, due to the importance of Cdt1, it is also regulated in other ways including the binding of the inhibitory protein geminin, which, in turn, is degraded in late mitosis and G1 phase (McGarry and Kirschner, 1998, Wohlschlegel et al., 2000). In addition, the interaction with chromatin-bound proliferating cell nuclear antigen (PCNA) directs Cdt1 for ubiquitination and subsequent degradation during S phase (Arias and Walter, 2006).

1.2.4.2 Regulation of origin firing by S-phase CDKs and Cdc7-Dbf4

In budding yeast, the accumulation of G1 cyclins leads to a gradual increase in CDK activity. The CDKs are then able to phosphorylate Whi5, an inhibitor of the transcription factors SBF and MBF, which in turn leads to the increased expression of a number of replication factors including S-phase cyclins (Nasmyth and Dirick, 1991, Spellman et al., 1998, Costanzo et al., 2004, Eser et al., 2011). CDKs also phosphorylate and promote the degradation of the CDK inhibitor Sic1, thus increasing the overall CDK activity in the cell, both through regulation of kinase expression and preventing the inhibition of CDKs (Feldman et al., 1997, Verma et al., 1997, Koivomagi et al., 2011). In S phase, the key role of the APC/C complex is also diminished. In G1, the APC/C complex binds to the adaptor protein Cdh1. In this active form, APC/C-Cdh1 is responsible for the degradation of Dbf4, the activation subunit of Cdc7 (Oshiro et al., 1999, Weinreich and Stillman, 1999, Ferreira et al., 2000). In S phase, the increase in CDK activity leads to

phosphorylation of Cdh1, which prevents binding to APC/C complex (Zachariae et al., 1998, Jaspersen et al., 1999). This allows for the accumulation of Dbf4 in the cell and formation of the catalytically competent Cdc7-Dbf4 heterodimer, which in concert with the CDKs promotes origin firing (alongside the role of the CDKs in preventing re-licensing of the duplicated genome).

The predominant role of CDK in budding yeast appears to be the phosphorylation of Sld2 and 3, facilitating Dpb11 binding, which is the minimal requirement for S phase entry and progression (Tanaka et al., 2007, Zegerman and Diffley, 2007, Kumagai et al., 2010, Kumagai et al., 2011). This in turn facilitates the transition of the pre-RC to the pre-LC. The phosphorylation of Sld3 by CDK is conserved in metazoan systems and the human homologue of Sld3 (Treslin), binds to the human homologue of Dpb11 (TOPBP1) in a CDK-dependent manner (Kumagai et al., 2010, Boos et al., 2011, Kumagai et al., 2011). However, the proposed human homolog of Sld2 (RecQL4) is able to bind to TopBP1 in the absence of CDK suggesting the minimal requirements of CDK activity vary across model systems.

The predominant function of Cdc7-Dbf4 is the phosphorylation of the N-terminal tails of the Mcm2-7 subunits, in particular Mcm2, 4 and 6 (Masai et al., 2006, Montagnoli et al., 2006, Sheu and Stillman, 2006, Randell et al., 2010). Phosphorylation target sites of Cdc7-Dbf4 have a limited amino acid sequence consensus, with the P+1 position occupied by a negatively charged residue (Asp, Glu or a phospho-Ser/Thr) (Cho et al., 2006, Charych et al., 2008). Phosphorylation of the Mcm2-7 complex is thought to induce an important conformational change in the helicase, which is required for activation. This is supported by the observation that removing residues 74-174 of Mcm4 allows for DNA replication *in vivo* in the absence of Cdc7-Dbf4 (Sheu and Stillman, 2010). This is suggestive of Cdc7-Dbf4 relieving an inhibitory action of the N-terminus of Mcm4 to promote origin firing. Another class of budding yeast mutants known to bypass the requirement for Cdc7-Dbf4 occurs within Mcm5 (Hardy et al., 1997). This is an interesting mutant, as Mcm5 has not been implicated as a potential target for Cdc7 activity (Lei et al., 1997). These data indicate that the structural changes triggered by Cdc7-Dbf4 may affect the global structure of Mcm2-7. However, structural studies of the phosphorylated Mcm2-7 showed little in the way of gross conformational

rearrangements within the hexamer (On et al., 2014). Activity of Cdc7-Dbf4 is also important for the recruitment of Sld3/7 and Cdc45 to origins of replication both *in vitro* and *in vivo* with multiple phosphorylation sites on the Mcm2-7 tails appearing to be integral to successful Sld3 loading (Yabuuchi et al., 2006, Heller et al., 2011, Tanaka et al., 2011, Deegan et al., 2016).

1.2.5 Key differences between yeast and human systems in origin loading and firing

The vast majority of studies on the mechanism and regulation of origin loading and firing for DNA replication have been performed in yeast model systems. This is due to the relative ease of performing such studies in yeast and the considerable functional conservation across eukaryotes. However, important differences exist between yeast and higher eukaryotes, which are important to consider when working in a human system.

The final product of origin licensing in both yeast and human consists of a loaded double hexamer suggesting that the mechanism of licensing is highly conserved across eukaryotes. To prevent premature firing of origins, CDK activity levels are kept low. Regulation of this process are also highly conserved, particularly the phosphorylation of and cyclin binding of Cdc6 to prevent relicensing outside of S-phase which both prevent origin binding and direct it for degradation by the the 26S proteasome (Drury et al., 2000, Mimura et al., 2004). In yeast, a further method to prevent relicensing is the exclusion of Mcms from the nucleus outside of G1 triggered by the CDK dependent phosphorylation of Mcm3 (Nguyen et al., 2000, Mimura et al., 2004). Importantly in many higher eukaryotes, exclusion of Mcms does not occur and in human cells Cdc6 is exported from the nucleus outside of G1 phase in a Crm1 dependent fashion upon phosphorylation by CDK2 (Jiang et al., 1999b). However, in contrast to the typical role of CDKs in regulation of licensing, phosphorylation of Cdc6 in human cells by CDK2-CyclinE has been shown to stabilise Cdc6 and facilitate origin licensing in cells that are exiting quiescence and beginning to proliferate. This phosphorylation event prevents association of Cdc6 with the APC/Cdh1 complex and allows accumulation of Cdc6 prior to the

accumulation of geminin and cyclin A as cells approach S-phase. This allows for efficient licensing prior to S-phase entry (Mailand and Diffley, 2005). This suggests an important interplay between the Cyclin A and Cyclin E bound forms of CDK2 in ensuring properly regulated DNA replication.

Perhaps the most significant difference between the Yeast and Human systems is the protein Geminin. Geminin is a regulatory protein present only in higher eukaryotes, which plays an essential role in preventing re-licensing of origins and subsequent re-replication during S-phase (Wohlschlegel et al., 2000). Binding of geminin to Cdt1 prevents loading of Mcm2-7 complexes and licensing is facilitated in G1 through the APC dependent degradation of geminin in late mitosis and G1 (McGarry and Kirschner, 1998). Cdt1 is a focal point for prevention of re-licensing in mammalian cells in which it is actively targeted for degradation outside of G1 which occurs through multiple distinct degradation pathways including the SCF-Skp2 E3 ubiquitin ligase complex in a CDK dependent fashion as well as well as the Cul4–Ddb1–Cdt2 E3 ubiquitin ligase complex which is directed by PCNA. This method of degradation also occurs in fission yeast (Gopalakrishnan et al., 2001, Hu and Xiong, 2006). However in budding yeast regulation of Cdt1 is achieved through its export from the nucleus along with the Mcm2-7 proteins during S-phase (Nguyen et al., 2000, Nguyen et al., 2001). The multiple redundant pathways centered upon regulating origin licensing through Cdt1 highlight its importance as a key regulator in DNA replication. This is further highlighted by the observation that over expression of Cdt1 or down regulation of geminin is sufficient to induce re-replication and subsequent genome instability. (Mihaylov et al., 2002, Arias and Walter, 2007).

Differences between yeast and higher eukaryotes in origin firing centre on the SDS complex (Sld2, Sld3 and Dpb11) for which functional homologues have been identified in higher eukaryotes including humans. As previously stated, Sld2 and Sld3 represent the minimal CDK targets for successful origin firing in yeast. Despite the conservation of Dpb11, Sld2 and Sld3 in humans (TopBP1, RecQL4 and Treslin), sequence conservation for these proteins is much less than for other replication factors and there are functional differences in their roles in origin firing. Treslin and TopBP1 have retained their crucial role in origin firing, with Treslin

showing significant homology to Sld3 and interacting with TopBP1 in a CDK dependent manner analogous to that seen in yeast (Sanchez-Pulido et al., 2010, Boos et al., 2011, Kumagai et al., 2011).

RecQL4, the functional homologue of Sld2, shares very limited sequence identity with its yeast counterpart and also appears to have diverged in terms of its binding to TopBP1. While the BCRT repeats required for RecQL4 binding (IV and V) are conserved in TopBP1, binding of RecQL4 has been shown to occur in a CDK independent manner (Makiniemi et al., 2001, Matsuno et al., 2006, Kumagai et al., 2010) and the phosphorylation sites present in Sld2 required for binding to Dpb11 are not conserved (Tanaka et al., 2007, Zegerman and Diffley, 2007). However much of this work has been conducted in *Xenopus* systems and needs to be confirmed in human cell lines. This is of particular importance, as *C. elegans* requires CDK activity for the recruitment of Sld2. This data suggests possible divergence in function even among higher eukaryotes (Gaggioli et al., 2014). In contrast to Sld2 comprising part of the pre-LC in yeast, RecQL4 has been shown to form a complex with the CMG, Ctf4 and Mcm10 (Xu et al., 2009, Im et al., 2015) suggesting that the role of RecQL4 has changed in evolution to fulfil a function in elongation. Studies in human cells have suggested that the N-terminal region of RecQL4 which shows homology with Sld2 is sufficient for the origin firing role of the protein (Kohzaki et al., 2012) suggesting the C-terminal helicase domain may play a role in elongation. This has yet to be confirmed but this region has at least been shown to be essential in *Drosophila.* The function of this region of RecQL4 requires significant further investigation in a number of metazoan systems. It is at least clear that RecQL4 is required for the formation of the CMG complex in human cells (Im et al., 2009) and that it may also play a role in the recruitment of polymerase α. The current data suggests that RecQL4 may bind to the replisome prior to S-phase to facilitate correct CMG formation before fulfilling its role in elongation after origin firing. This is in stark contrast to the yeast system in which Sld2 is released upon DNA unwinding.

Further to these differences, two more proteins have been identified in higher eukaryotes, which bind to TopBP1. Geminin coiled coil domain containing 1 (GEMC1) appears to have a role in facilitating Cdc45 recruitment. It is also heavily

phosphorylated by CDK2-Cyclin E and this activity contributes to its ability to stimulate DNA replication (Balestrini et al., 2010). DNA unwinding element binding protein B (DUE-B) has also been shown to associate with Cdc45 (Chowdhury et al., 2010) and its ability to form complexes with the Mcm2-7 complex is regulated by phosphorylation of its C-terminus. This phosphorylation is regulated by the interplay between Cdc7 and PP2A activity, suggesting a novel regulatory mechanism for Cdc45 recruitment in higher eukaryotes (Gao et al., 2014).

It is clear from this data that there is still much left to discover about the initiation of DNA replication in humans and other higher eukaryotes. While yeast has provided a powerful tool as a model system, when studying replication in human cells it is important to appreciate that many factors may exist that have not yet been identified.

1.3 Cdc7-Dbf4

Cdc7 and its activating subunit Dbf4 were first identified in budding yeast genetic screens. Cdc7 mutants lead to an S-phase arrest in non-permissive conditions (Hartwell, 1971, Hartwell, 1973). The arrest of cells with 1C DNA content later confirmed the importance of Cdc7 activity in the initiation of DNA replication (Hereford and Hartwell, 1974). In the following years, the *CDC7* gene was confirmed as a Ser/Thr kinase and its importance in both mitosis and meiosis was reported (Patterson et al., 1986, Bahman et al., 1988, Hollingsworth and Sclafani, 1990, Buck et al., 1991b, Yoon and Campbell, 1991). Dbf4 mutants also lead to an S-phase arrest phenotype, leading to cells that formed a dumbbell-like shape which is often associated with a defect in the initiation of DNA replication (Johnston and Thomas, 1982b, Johnston and Thomas, 1982a). The similar phenotypes of Cdc7 and Dbf4 suggested that they both played a role in the G1/S-phase transition. The interaction between the two was first suggested due to the synthetic lethality of their temperature sensitive mutants (Kitada et al., 1992). Overexpression of Dbf4 was also able to rescue the temperature sensitive mutant of Cdc7 suggesting a direct interaction between Cdc7 and Dbf4, where the latter modulated the activity of the former (Kitada et al., 1992).

1.3.1 Regulation of Cdc7-Dbf4

Cdc7-Dbf4 activity peaks at the G1/S-phase boundary and is maintained at high levels throughout S-phase (Jackson et al., 1993, Oshiro et al., 1999). Dbf4 protein levels are shown to oscillate throughout the cell cycle with the highest levels of expression coinciding with a peak in kinase activity, while Cdc7 levels remain relatively stable throughout the cell cycle (Oshiro et al., 1999, Weinreich and Stillman, 1999, Ferreira et al., 2000). Total levels of Dbf4 in the cell also correlated with the amount of Dbf4 immunoprecipitated with Cdc7 in the above studies.

Dbf4 levels are regulated by the APC/C, which targets it for degradation except in S phase, when the activity of APC/C is inhibited (Oshiro et al., 1999, Weinreich and Stillman, 1999, Ferreira et al., 2000). In eukaryotes, replication origins can be classified by the time they fire as early, intermediate and late and these relative firing times appear to be decided in G1 (Dimitrova and Gilbert, 1999). Concordantly, Cdc7-Dbf4 is essential not only for G1/S transition, but also for the continued progression of S phase to its completion (Bousset and Diffley, 1998, Donaldson et al., 1998). Moreover, recent evidence indicates that Cdc7 play a role in the temporal regulation of S phase progression and efficiency of firing of individual origins. This activity is regulated by the telomere binding protein Rap-interacting factor 1 (Rif1). Depletion of Rif1 leads to a loss of mid-S replication foci profiles, reduced stimulation of initiation events in early S-phase and changes in long range replication timing domain structures (Yamazaki et al., 2012). This is due to Rif1 targeting protein phosphatase 1 (PP1) to replication origins and counteracting Cdc7-Dbf4 activity. Rif1 is itself also regulated by Cdc7-Dbf4 phosphorylation, which creates a binding site for PP1. This system allows for timely replication across all of the replication origins and this timing is tightly linked to Cdc7-Dbf4 activity (Dave et al., 2014, Hiraga et al., 2014, Mattarocci et al., 2014). The fission yeast ortholog of Cdc7-Dbf4, Hsk1-Dfp1, was also shown to regulate the timing of origin firing, with higher local levels of the kinase increasing the odds of firing for less efficient origins (Patel et al., 2008)

Cdc7 and its activation subunit are conserved from yeast through to humans, with orthologs identified in fission yeast and a number of metazoan species. The human version of Cdc7 was discovered by 1997, and its partner Dbf4, also known as the activator of the S-phase kinase (ASK), was identified in yeast two-hybrid screens two years later (Sato et al., 1997, Jiang et al., 1999a, Kumagai et al., 1999). Drf1, the alternative Cdc7 activator, was identified more recently in a search for proteins that shared sequence homology with Dbf4 (Montagnoli et al., 2002). It was then shown that Drf1 interacted with and activated Cdc7 *in vitro*. Drf1 has so far only been found in metazoans, and its functions have not been well studied. Some evidence suggests that it may play roles in early development in *Xenopus* (Takahashi and Walter, 2005a) as well as fulfilling the many roles of Dbf4 in Xenopus extracts including DNA replication, checkpoint responses and coordinating chromosome cohesion with replication (Takahashi and Walter, 2005a, Silva et al., 2006, Takahashi et al., 2008, Tsuji et al., 2008). When the cell acquires somatic characteristics, Drf1 is depleted and replaced by Dbf4 for the cellular functions of Cdc7 (Takahashi and Walter, 2005b)**.** Human Cdc7 and Dbf4 exhibit similar patterns of activity and expression as their budding yeast counterparts. Transcription of the human *DBF4* gene and levels of the protein oscillate with mRNA levels at their lowest at the G2/M phase of the cycle and steadily increasing towards entry into S phase. Dbf4 protein levels also peak during S phase with levels starting to increase from late in G1 (Kumagai et al., 1999). As expected the human Cdc7-Dbf4 kinase activity strongly correlated with expression of Dbf4 (Kumagai et al., 1999). These results must be taken with the caveat that regulation of Drf1 remains largely unexplored. Conditional knockout of Cdc7 in mouse embryonic stem (ES) cells also lead to cells being blocked in S-phase with a reduction in DNA synthesis (Kim et al., 1998) and in a mouse model, deletion of *Cdc7* was embryonically lethal (Kim et al., 2002).This collective data not only highlights the importance of Cdc7 activity in a number of systems, but also the functional conservation of the activation of Cdc7-Dbf4 across all eukaryotes. The conservation of Cdc7-Dbf4 activation and action is further demonstrated by the ability of synthetic human *CDC7* and *DBF4* genes to support DNA replication and cell division in a budding yeast system (Davey et al., 2011).

1.3.2 Key differences in Cdc7 regulation in yeast and human systems

While this functional conservation highlights the importance of Cdc7 activity and its correct regulation it is important to appreciate there are differences in the regulation of Cdc7 activity between yeast and higher eukaryotes such as in human systems. As previously discussed, regulation of Cdc7 in both systems is largely based on binding of Dbf4, the levels of which peak at the G1/S phase transition and throughout S-phase. In both systems levels of Dbf4 are regulated through APC mediated degradation. A key difference between model systems is the chromatin binding of Cdc7 and Dbf4. In budding yeast, Cdc7 remains associated with chromatin throughout the cell cycle while Dbf4 binds at the G1/S-phase transition in an ORC dependent fashion and remains throughout S-phase (Weinreich and Stillman, 1999, Duncker et al., 2002). In the human system, ASK is accumulates in the nucleus at the post mitotic phase but does not associate to with chromatin until late G1. This is also the case when ASK is ectopically expressed. Cdc7 however binds to chromatin immediately after mitosis and remains bound through G1 and S-phase (Sato et al., 2003). This spatial regulation of the kinase and its activating subunit will have implications for origin firing across different model systems. Further to the binding of Dbf4, recent data has revealed a role for phosphorylation of Cdc7 as a means of regulating its activity outside of S-phase. Phosphorylation of a number of sites on Cdc7 by CDK1 leads to the dissociation of Cdc7 from origins. This prevents inappropriate replication occurring in mitosis and possible genome instability. Protein phosphatase 1 α (PP1α) is then required to dephosphorylate Cdc7 to facilitate the next round of replication (Knockleby et al., 2016). A recent study in mammalian cells has also revealed a method of regulation in which p53 influences Cdc7 levels in response to genotoxic stress to enforce a block in the G1 phase. This occurs both post-transcriptionally to reduce Cdc7 expression but also through ubiquitination to promote its degradation. Such mechanisms of Cdc7 regulation have yet to be described in yeast (Tudzarova et al., 2016).

Phosphorylation of Cdc7-Dbf4 as a means of regulating its activity in response to genotoxic stress is discussed in more detail in section 1.4.1. Lastly, the most significant difference between yeast and human systems is the presence of an alternative activator of Cdc7 in higher eukaryotes called Drf1 (Montagnoli et al.,

2002). Drf1 has been implicated in a number of cellular functions and is discussed in detail in section 1.4.5 as well as in the relevant sections about the various roles of Cdc7 in the cell.

1.3.3 Structural organisation of Cdc7-Dbf4

1.3.3.1 Structural and functional features of Cdc7 identified by early studies

Cdc7 contains a number of conserved kinase motifs (I-XI), which are common to all Ser/Thr kinases. However, the canonical kinase organisation is interrupted by kinase insert (KI) sequences. In budding yeast, three such inserts (named KI1, KI2 and KI3) can be identified, while only two (KI2 and KI3) are present in metazoan orthologs. The position of the insert sequences relative to the conserved kinase domains is the same across all eukaryotes with KI1 (when present) between I and II, KI2 between VII and VIII and KI3 between X and XI (Patterson et al., 1986, Hanks et al., 1988, Masai et al., 1995, Jiang and Hunter, 1997, Sato et al., 1997, Faul et al., 1999, Guo and Lee, 1999). However, the lengths of the kinase inserts and their amino acid sequence varies widely among Cdc7 orthologs, and they are predicted to be largely devoid of secondary structure. Differences between closely related orthologs of Cdc7 can also be seen as with budding and fission yeast (Figure 1-8). Despite the two proteins being identical in length budding yeast Cdc7 has a significantly larger KI2 and KI3 than its fission yeast counterpart. However, the fission yeast ortholog appears to compensate for this with a larger KI1 and significantly longer N- and C-terminal non-conserved regions. This suggests that the structural feature required for activity of each kinase may vary significantly and these differences are likely to be even greater in metazoan species. Of note, the presence of kinase insert sequences as such, is not unique to Cdc7. For instance, protein kinase R (PKR), a kinase that phosphorylates eukaryotic initiation factor 2alpha (eIF2α) to regulate the initiation of translation contains an insert sequence between catalytic sub domains IV and V. Deletion of this insert as well as point mutation of Ser355 abrogates kinase activity while having no effect on substrate binding *in vivo* (Craig et al., 1996). The mitogen activated protein kinases (MAPKs) also contain an insert sequence between the αG and αH helices of the C-lobe (Canagarajah et al., 1997). In the case of the MAPK Erk2 (extracellular signal-

related kinase 2), this insert becomes disordered upon phosphorylation of the kinase to allow for kinase activation through decreased interaction with the phosphorylation lip of Erk2 as well as having a role in substrate binding (Canagarajah et al., 1997, Chou et al., 2003). An interesting example of a poorly conserved kinase insert sequence is that of Greatwall kinase (Gwl). Gwl is a kinase required for M phase entry and maintenance across a number of model systems (Yu et al., 2004, Burgess et al., 2010). Greatwall specifically deactivates protein phosphatase 2A (PP2A) bound to B55 type regulatory subunits to prevent the removal of M-phase specific phosphorylation events catalyzed by CDK1-CyclinB (Vigneron et al., 2009). Greatwall kinase contains a particularly large insert sequence of approximately 500 amino acids termed the non-conserved middle region (NCMR), which splits the conserved N and C-lobes of the kinase. Deletions in the NCMR have minimal affect on kinase activity *in vitro* and the ability of the kinase to maintain its mitotic functions in the *Xenopus* system (Blake-Hodek et al., 2012). Much like Cdc7, structural information about Gwl is largely restricted to conserved kinase sequences. It has also recently attracted attention as a possible target for ant-cancer therapeutics and as such represents another important but challenging kinase for structural studies (Ocasio et al., 2016).

Extensive mutagenesis of human Cdc7 undertaken showed that large portions of the insert sequences are not strictly essential for the specific enzymatic activity *in vitro* (Kitamura et al., 2011, Hughes et al., 2012). In particular, deletions in KI3 required for the initial crystallization had no effect on the ability of the kinase to phosphorylate an Mcm2-derived substrate peptide in *vitro*. However, a deletion of 132 residues (Δ228-359) in KI2 did have a deleterious effect on activity. However, it is important to note that KI2 is present between the conserved DFG and APE motifs of Cdc7 and thus extends the activation loop of Cdc7 by 172 amino acids (Hughes et al., 2012). Kinase inserts 2 and 3 were reported to play a role in the nuclear import and export of Cdc7 (Kim and Lee, 2006, Kim et al., 2007). Thus, a nuclear localisation signal (NLS), identified within KI2, mediates the interaction with importin-β and facilitates nuclear import of the kinase (Kim and Lee, 2006). Moreover, a nuclear retention (NRS) sequence present within the NLS is essential for chromatin binding upon entry into the nucleus. One of two nuclear export

sequences (NES) is present in KI3, highlighting the role of the insert sequences in the localisation and spatial regulation of Cdc7 (Kim et al., 2007).

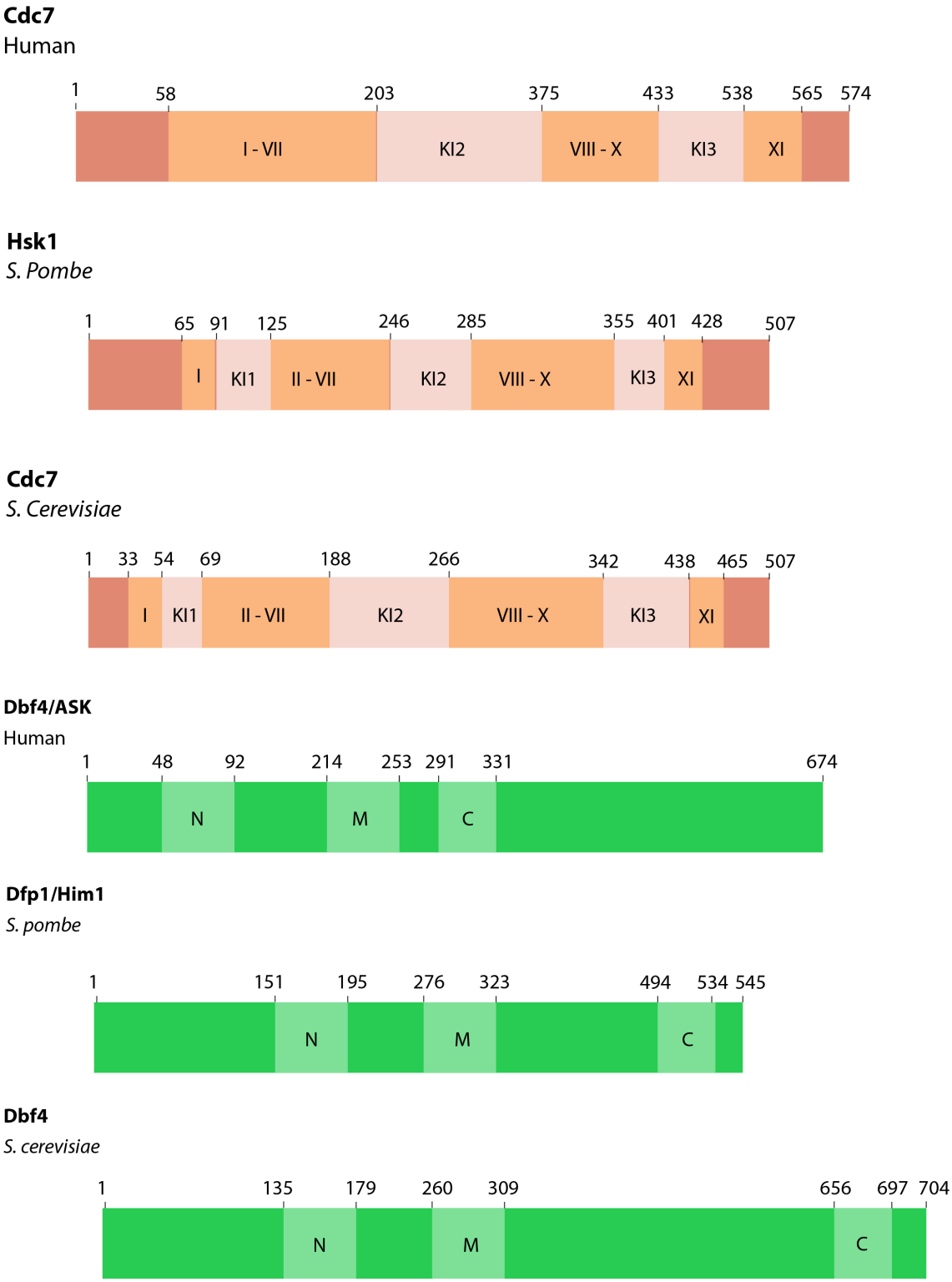


Figure 1-8 Schematic of the *S. cerevisiae* and human Cdc7 and Dbf4 structures

Conserved kinase motifs of Cdc7 are labelled with roman numerals (I-XI) and insert sequences labelled as KI1, KI2 and KI3. Non-labelled portions of Cdc7 are non-conserved N-terminal regions. The conserved motifs-N, M and C of Dbf4 are also indicated.

1.3.3.2 Domain organisation of Dbf4

Across orthologs of Dbf4, the amino acid sequence conservation and predicted secondary structures are confined to three short regions named by the their relative positions along the protein sequence: motif-N, -M and -C (Kumagai et al., 1999, Takeda et al., 1999, Masai and Arai, 2000). The relative positions of these motifs along the protein vary amongst Dbf4 orthologues (Fig 1-8). The N-motif of Dbf4 in human is considerably closer to the N-terminus than is seen in budding and fission yeast, and the contains a significant C-terminal sequence of ~300 amino acids which appears to be entirely devoid of conserved features. In the case of the two yeast orthologues, the C-motif is present towards the extreme C-terminus of the protein with a an extended stretch between the M and C motifs relative to that seen in human. Interestingly between budding yeast Dbf4 and Fission yeast Dfp1 there is a significant difference in the protein size that appears to be attributed to a significant decrease in the number of residues between the M and C motifs. This suggests that in fission yeast at least, a large stretch of this region of the protein was not essential for cell survival, while in budding yeast it may play an important role in the cell. These domains appear to act as Cdc7 binding/activation and chromatin tethering modules. A small fragment spanning motifs -M and -C of Dbf4 is sufficient for Cdc7 binding and activation *in vitro* (Ogino et al., 2001, Ogino et al., 2006, Kitamura et al., 2011, Hughes et al., 2012). The segment connecting the two motifs is poorly conserved and variable in length, although a string of positively charged residues is important for the kinase activity *in vitro* (Hughes et al., 2012). The importance of motifs-M and -C was clearly demonstrated in the yeast systems. In one study, deletion mutants of Dfp1 were used to complement the growth defect of *S. pombe dfp1-null* strains. Deletions of the N-terminal 222 amino acids did not affect the ability of the construct to complement the growth defect of the *Dfp1*-null cells. However further deletion up to the position 335 lead to a loss of complementation due to the removal of motif-M. Removing as few as 35 residues at the C-terminal end of Dfp1 also prevented complementation due to the disruption of the C-motif. This confirmed the importance of motifs -M and -C for Hsk1-Dfp1 activity, while suggesting that motif-N of Dfp1 is dispensable under some conditions *in vivo*. Furthermore, small segments of the protein spanning motifs-M and -C could

bind Hsk1 individually, with binding of both being required for the activation of the kinase (Ogino et al., 2001). Similar observations were also made in budding yeast, where a fragment of Dbf4 spanning motif-M and the 100 amino acids immediately downstream was able to bind Cdc7, as was motif-C (Harkins et al., 2009). In humans, a minimal fragment of Dbf4 spanning motifs-M and C (174-350) was also capable of supporting kinase activity *in vitro*; however, addition of more residues preceding the motif-M (134-350) was needed for hyper-phosphorylation of Mcm2 *in vivo.*

Motif-N of Dbf4 is dispensable for the kinase activity *in vitro* as well as for its ability to support the cell cycle in a number of systems. However, it has been suggested to play a role in the interaction with the replication machinery. Dbf4 co-localises with Orc2 and co-fractionates with it in the chromatin insoluble fraction in a manner dependent on correct formation of the ORC complex in yeast (Pasero et al., 1999). A direct interaction between Dbf4 and the Orc2 and 3 subunits was corroborated by later studies, and was shown to be reliant on motif-N (Duncker et al., 2002, Varrin et al., 2005). An N-terminal fragment of Dbf4 (110-296) was also shown to be required for the interaction with the loaded helicase of the pre-RC (Francis et al., 2009). In the *Xenopus* system, not only Dbf4 activates Cdc7, but it also recruits it to the pre-RC. Dbf4 binds chromatin early in the cell cycle independently of origin licensing. Cdc7 is then recruited to chromatin upon Mcm2-7 loading. Immunodepletion of Dbf4 prevents this recruitment, indicating that the activating subunit acts as the chromatin-tethering module (Jares and Blow, 2000, Jares et al., 2004). It is remarkable that a protein that is so evolutionarily diverse can maintain such an essential function. The presence of discreet modules in Dbf4 is likely necessary to allow Dbf4 to interact with a range of different proteins and fulfil multiple roles without the need for the expression of a number of different genes. This idea is further supported by the interaction between Dbf4 and a number of kinases implicated in cell cycle regulation, such as Rad53 and Cdc5 (Chen and Weinreich, 2010, Matthews et al., 2012). Very little is known about the function of Dbf4 regions outside of the three motifs, which collectively comprise over 80% of the protein's mass.

1.4 Other roles of Cdc7-Dbf4

Aside from its function in the initiation of DNA replication, Cdc7 plays a number of additional critical roles in the cell. Notably, Cdc7 is involved in mitotic exit, chromosome cohesion, responses to replication stress, as well as participating in DNA recombination and chromosome segregation in meiosis. Further to this, Cdc7 has an alternative binding partner called Drf1, which functions alongside Cdc7 during development.

1.4.1 The role of Cdc7 in replication stress

In addition to its well-established role during the normal cell cycle, Cdc7 has been implicated in responses to replication stress and, in particular, in the intra S-phase checkpoint. Aberrant DNA replication during S phase triggers suppression of late origin firing and stabilisation of replication forks. This process decreases the amount of damage caused by collapsed forks, while helping to salvage replication after the block has been removed (Segurado and Tercero, 2009).

The evidence of the involvement of Cdc7 in response to replication stress came from early studies in budding yeast. *S. cerevisiae* strains lacking Cdc7 and capable of growth in the presence of the suppressing *mcm5-bob1* mutation are hypersensitive to the treatment with hydroxyurea (HU) (Weinreich and Stillman, 1999). Exposure to HU leads to hyper-phosphorylation of Dbf4 by the checkpoint kinase Rad53 (Chk2 in humans). Interestingly, Cdc7 also phosphorylates Rad53 in a manner that is important for full Rad53 activity upon checkpoint activation (Dohrmann et al., 1999, Weinreich and Stillman, 1999, Kihara et al., 2000). Furthermore, Rad53 was shown to interact with motif-N of Dbf4 and a deletion of this motif leads to a defect in the checkpoint response upon HU treatment (Ogino et al., 2001, Duncker et al., 2002, Varrin et al., 2005). Importantly, Dbf4 motif-N was shown to mediate an interaction with ORC (Pasero et al., 1999). Therefore, phosphorylation by Rad53 represents a method of preventing Cdc7 association with ORC following DNA damage, displacing it from replication origins (Santocanale and Diffley, 1998, Weinreich and Stillman, 1999, Duncker et al., 2002,

Varrin et al., 2005). A similar relationship exists between Hsk1 and the checkpoint kinase Cds1 in fission yeast (the homologue of Rad53), where the deletion of Dfp1 motif-N also leads to defects in S-phase checkpoint control (Takeda et al., 1999). Cds1 hyper-phosphorylates Dfp1 upon HU treatment and subsequent replication fork stalling. Cds1 also phosphorylates Hsk1, and Hsk1 activity is required for full activity of Cds1 during HU treatment. (Brown and Kelly, 1999, Snaith et al., 2000, Takeda et al., 2001). The restart of replication and subsequent completion of mitosis upon HU treatment also depends on Hsk1 (Snaith et al., 2000).

The importance of Cdc7 in checkpoint responses is maintained in higher eukaryotes. A decrease in phosphorylation of the checkpoint kinase Chk1 is observed in Cdc7-depleted mammalian cells upon HU treatment, increasing the sensitivity of cells to such treatment. A decrease in Chk1 activation was also seen in *CDC7*-null mouse ES cells, when subject to HU or UV treatment (Kim et al., 2008). Depletion of Cdc7 also leads to a decrease of claspin phosphorylation and its association with chromatin. Claspin is an essential mediator of replication stress checkpoint that plays a critical role in the regulation of Chk1 (Chini and Chen, 2003, Chini and Chen, 2004, Kim et al., 2008). Conditional knockout of claspin in MEFs leads to S-phase defects due to a role in facilitating normal DNA replication. Cdc7 binds an acidic patch in claspin. This interaction triggers phosphorylation of claspin by Cdc7 and prevents an intramolecular interaction between the acidic patch of claspin and its DNA binding domain. This data suggests a role for claspin in initiation of DNA replication through recruitment of Cdc7, in addition to its role in the S-phase checkpoint (Yang et al., 2016). In human cells, Cdc7 is inhibited by ATR-dependent prevention of Dbf4 binding (Costanzo et al., 2003, Dierov et al., 2004). However, other data suggests that Cdc7-Dbf4 activity is not decreased during genotoxic treatments (Dierov et al., 2004, Tenca et al., 2007, Tsuji et al., 2008). Phosphorylation of Dbf4 upon genotoxic stress by checkpoint kinases was thought to be an integral part of the S-phase checkpoint, which ultimately leads to a decrease in origin firing (Lee et al., 2012). However, the same study concluded that phosphorylation of Dbf4 had little effect on the activity of Cdc7 and was likely involved in mediating a protein-protein interaction (Lee et al., 2012).

Cdc7-Dbf4 was shown to accumulate on chromatin in response to genotoxic stress, which is facilitated by ATR-Chk1 dependent inhibition of the APC/C^{Cdh1} complex that is responsible for Dbf4 degradation (Yamada et al., 2013). It is thought that activation of the ATR-Chk1 pathway leads to auto-degradation of Cdh1, which in turns stabilises Cdc7-Dbf4 to secure replication restart after the block has been removed (Yamada et al., 2013). It is thought that stabilisation of Cdc7-Dbf4 is required for deciding which DNA repair pathway to utilise prior to replication restart. Cdc7-Dbf4 was shown to have a role in translesion synthesis in yeast (Njagi and Kilbey, 1982, Pessoa-Brandao and Sclafani, 2004). Furthermore, recent data in human cells demonstrated Cdc7-Dbf4 - dependent phosphorylation of RAD18 that leads to its interaction with the translesion synthesis (TLS) DNA polymerase η (Pol η) (Day et al., 2010, Vaziri and Masai, 2010). RAD18 also interacts with the Zn-binding motif of Dbf4, therefore the translesion DNA synthesis machinery is recruited through Dbf4 via RAD18 (Yamada et al., 2013). This is an appealing concept for decisions regarding DNA repair pathways, as RAD18 is also known to interact with Rad51C, an essential factor for homologous recombination (Huang et al., 2009).

The maintenance of Cdc7-Dbf4 on chromatin and its continued activity seem at odds with the suppression of late origin firing. A study in *Xenopus* suggested a possible role for PP1, which can be activated by checkpoint kinases to reverse phosphorylation of the Mcm2-7 complex, without affecting the activity of Cdc7-Dbf4 itself (Poh et al., 2014). However, there is no evidence as yet that the depletion of such phosphatases prevents the inhibition of late origin firing. It is clear that the role of Cdc7 under genotoxic stress is complex and multifaceted, and this is likely to represent a large portion of future research into the many functions of this kinase.

1.4.2 The role of Cdc7 in Meiosis

Meiosis is a specialised cell division programme used for the production of haploid gametes. In meiosis, two consecutive chromosome segregation events occur after a single round of DNA replication (Sakuno and Watanabe, 2009). As with mitotic cells, this DNA replication requires Cdc7, however the kinase also plays a further

role in the initial chromosome segregation event (meiosis I). During this stage, Cdc7 is required for the initiation of meiotic recombination and chromosome segregation (Buck et al., 1991a, Ogino et al., 2006, Sasanuma et al., 2008, Wan et al., 2008). Meiotic recombination involves the intentional formation of double-strand breaks (DSBs) in chromosomal DNA. Cdc7-Dbf4 was shown to promote the formations of DSBs in budding yeast through phosphorylation of Mer2, the accessory factor to the meiotic endonuclease Spo11 responsible for the DSB formation. Mer2 is phosphorylated by both Cdc7-Dbf4 and S-CDKs, and this phosphorylation is essential for the recruitment of Spo11 to DSB sites (Sasanuma et al., 2008, Wan et al., 2008).

Due to the *mcm5-bob1* mutation, which bypasses the need for Cdc7 in DNA replication initiation in yeast, it was possible to analyse the effect of Cdc7 depletion on meiosis. These studies revealed that Cdc7-Dbf4 is required for homologous chromosome segregation in meiosis I (Valentin et al., 2006, Lo et al., 2008, Matos et al., 2008). A failure to segregate in meiosis I leads to sister chromatid segregation in meiosis II resulting in a pair of diploid spores in a process similar to mitosis (Lo et al., 2008, Matos et al., 2008). Components of the replisome in yeast (Tof1 and Csm3) were also shown to recruit Cdc7-Dbf4 to the replisome, where it phosphorylates Mer2 in the wake of the replication fork. This coordinates DNA replication with the subsequent formation of DSBs for meiotic recombination (Murakami and Keeney, 2014). Cdc7-Dbf4 also plays a key role as a gene specific regulator of the global transcription factor NDT80, which plays an important role in the transition to meiotic division in fission yeast (Lo et al., 2012). Cdc7-Dbf4 increased NDT80 transcription by relieving repression mediated by a complex of Sum1 (suppressor of mar1-1 protein), Rfm1 (repression factor of MSEs protein 1), and the histone deacetylase; Hst1. Cdc7-Dbf4 phosphorylates Sum1 at 11 meiosis-specific sites. Sum1 is also phosphorylated by CDK1 and the meiosis-specific kinase Ime2, and this is a key regulatory step in the transition to the meiotic cell cycle.

Cdc7-Dbf4 also plays a more direct role in facilitating the separation of chromatids during meiosis I. Phosphorylation is a key regulatory step in determining whether cohesin is cleaved by separase during cell division, and the phosphatase PP2A

inhibits this cleavage event. Conversely, in concert with casein kinase 1 (CK1), Cdc7-Dbf4 promotes the cleavage of the cohesin subunit Rec8 and allows for chromosome segregation to proceed (Katis et al., 2010).

1.4.3 The role of Cdc7 in chromosome cohesion

Cdc7 was shown in a number of model systems to play a role in the establishment and maintenance of chromosome cohesion. A key aspect of chromosome cohesion is the loading of the cohesin complex by the Scc2-Scc4 complex during the G1 phase of the cell cycle (Uhlmann and Nasmyth, 1998, Lengronne et al., 2006). Cohesin loading is also dependent on the presence of the pre-RC in *Xenopus* egg extracts, and Cdc7 activity is required for the recruitment of Scc2-Scc4 to the pre-RC. Conversely, Cdc7 depletion significantly reduces the association of Scc2-Scc4 with chromatin (Gillespie and Hirano, 2004, Takahashi et al., 2004, Takahashi et al., 2008).

In budding yeast, the activity of Cdc7-Dbf4 was also shown to be important for the coordination of DNA replication and chromosome cohesion. Thus, Cdc7-Dbf4 accumulates on kinetochores in telophase, through an interaction with the Ctf19 kinetochore complex. This in turn allows for Sld3-Sld7 recruitment to pericentromeric replication origins, so that they initiate replication early in S phase. Cdc7-Dbf4 can then facilitate the recruitment of the Scc2-Scc4 cohesin loading complex, which facilitates robust sister chromatid cohesion at microtubule attachment sites (Natsume et al., 2013).

Similarly, in fission yeast, Hsk1-Dfp1 directly interacts with Swi6, the *S. pombe* ortholog of heterochromatin protein 1 (HP1). Swi6 is required for cohesin recruitment to heterochromatin and for establishing cohesion at centromeres. Hsk1-Dfp1 is able to phosphorylate Swi6 both *in vitro* and *in vivo* (Bernard et al., 2001, Nonaka et al., 2002, Bailis et al., 2003b). Truncation mutant of Dfp1 containing only the first 376 residues of the protein showed a drastic reduction in Rad21, an essential component of the cohesin complex, at centromeric heterochromatin but not along the arms of the chromosome. This phenotype is

typical of *swi6*-null cells, highlighting the importance of the interaction between the Hsk1-Dfp1 and Swi6 for proper maintenance of chromosome cohesion (Bailis et al., 2003a).

1.4.4 The role of Cdc7 in mitotic exit

Cdc7 was also shown to interact with Cdc5, the single Polo kinase in budding yeast. Cdc5 plays an important role in regulating mitotic progression and cytokinesis, including its recently elucidated role in facilitating the removal of centromeric cohesin during mitosis (Mishra et al., 2016). An interaction between the N-terminus of Dbf4 and a polo box domain (PBD) of Cdc5 alters the substrate targeting of Cdc5, preventing it from activating the mitotic exit network (MEN), without affecting the kinase activity of Cdc5 (Miller et al., 2009). Dbf4 mutants that are unable to interact with Cdc5 can support the progress through S-phase, but this is followed by aberrant chromosome segregation in mitosis. Dbf4 therefore acts as an inhibitor of Cdc5 to ensure cells exit mitosis correctly (Miller et al., 2009). The interaction between Dbf4 and the PBD of Cdc5 is unusual in that it does not require Dbf4 to be phosphorylated. Dbf4 contains an atypical PBD-interaction domain that binds the PBD via a region distinct from that used by phosphoproteins (Chen and Weinreich, 2010). Thus, Cdc7-Dbf4 is an important regulator in mitosis that prevents mis-segregation of chromosomes and subsequent genome instability.

1.4.5 The role of Cdc7-Drf1

Drf1 was initially identified through sequence similarity with both the yeast and human homologues of Dbf4 and was shown to co-purify with and to activate Cdc7 allowing for phosphorylation of an Mcm2-derived substrate (Montagnoli et al., 2002). Drf1 and Dbf4 do not bind Cdc7 simultaneously, suggesting these paralogs evolved to regulate Cdc7 for different functions in the cell (Montagnoli et al., 2002). However, expression levels of Drf1 are similar to that of Dbf4 in primary human fibroblasts throughout the cell cycle, peaking in S-phase and decreasing after mitosis suggesting some redundancy in function (Montagnoli et al., 2002).

In the *Xenopus* system, Drf1 was reported to play a role in the initiation of DNA replication during oogenesis and the early embryogenesis, before being replaced by Dbf4 after maturation. Thus, at least in frogs, Drf1 appears to be a developmentally regulated functional equivalent of Dbf4 (Takahashi and Walter, 2005a, Silva et al., 2006). Immunodepletion of both Drf1 and Dbf4 showed that they are found individually bound to Cdc7, and there is a 5-fold molar excess of Drf1 relative to Dbf4. Depletion of Dbf4 also had minimal effect on DNA replication, while depletion of Drf1 lead to significant inhibition of Mcm4 phosphorylation and subsequent DNA replication (Takahashi and Walter, 2005a, Silva et al., 2006). Cdc7-Drf1 is also the active form of the kinase required for regulation of chromosome cohesion in the *Xenopus* system (Takahashi et al., 2008).

1.5 Cdc7-Dbf4 and cancer

Due to its numerous roles in cell cycle progression and responses to replicative stress, Cdc7-Dbf4 has attracted a lot of attention as an attractive target for the development of cancer therapeutics (Sawa and Masai, 2009, Swords et al., 2010). Cdc7 is overexpressed in a number of cancers and this expression often correlates with patient prognosis (Nambiar et al., 2007, Bonte et al., 2008, Kaufmann et al., 2008, Clarke et al., 2009, Kulkarni et al., 2009, Hou et al., 2012, Cheng et al., 2013, Melling et al., 2015, Ghatalia et al., 2016). Knockdown of Cdc7 using siRNAs in cancer cell lines also leads to cell death through the accumulation of nuclear damage in S-phase, which is followed by apoptosis or mitotic crisis (Montagnoli et al., 2004, Im and Lee, 2008). Normal fibroblasts were able to survive such treatments, in large part due to checkpoint mechanisms regulated by p53, which is often mutated in cancers. Use of the cell cycle dye Fucci in cancer cell lines revealed that cell cycle defects caused by Cdc7 depletion vary greatly between p53-positive and negative cells. Cells negative for p53 arrest predominantly in G2 and accumulate proteins involved in mitotic regulation, such as cyclin B1. This leads to aberrant entry into mitosis and subsequent cell death. Cells positive for p53 do not accumulate in G2 and appear to die due to an aberrant S-phase. This knowledge of distinct mechanisms of cell death in different cellular contexts may allow for the development of treatments that could be used in conjunction with

Cdc7 inhibitors to achieve synergistic effects (Ito et al., 2012). Tumours expressing mutants of checkpoint-related genes are also likely to be hyper-sensitive to Cdc7 inhibitors, possibly allowing killing of cancerous cells, while sparing normal tissue.

Considerable efforts have already been expended to create Cdc7 inhibitors, with the first small molecules, heteroaryl-pyrrolopridinones, reported 8 years ago by Nerviano Medical Sciences (Vanotti et al., 2008). Optimisation of these compounds led to the development of PHA-767491, which inhibited Cdc7-Dbf4 - specific phosphorylation of Mcm2 and prevented origin firing without impeding replication fork progression (Montagnoli et al., 2008). Further to this, treatment of cancer cell lines with PHA-7674891 lead to apoptotic cell death, and tumour growth was inhibited in pre-clinical models. PHA-7674891 has shown some promise as a possible inhibitor for the treatment of pancreatic adenocarcinoma (Huggett et al., 2016), and was shown to have a synergistic effect in inhibiting both Cdc7 and CDK9 in hepatocarcinoma and to enhance the efficacy of 5-fluorouracil (Li et al., 2015b). In the following years, a more potent and specific Cdc7-Dbf4 inhibitor, XL413, was reported by Exelxis (Koltun et al., 2012), which advanced into clinical trials. However, the ability of XL413 to inhibit Cdc7 in various cancer cell lines was hampered by poor cell permeability, and work is now underway to identify new Cdc7 inhibitors that are more readily taken up by the cell (Sasi et al., 2014). Cdc7 inhibitors clearly represent a promising class of anti-cancer therapeutics and detailed structural and functional information about Cdc7-Dbf4 could inform the design of the next generation of inhibitors.

1.6 Crystal structure of a minimally active Cdc7-Dbf4 construct

Due to the inherent flexibility of Dbf4 and the insert sequences in Cdc7, structural characterisation of the heterodimeric kinase was a daunting challenge. Cdc7 displays no detectable kinase activity on its own (Ogino et al., 2001, Kitamura et al., 2011) and is prone to aggregation, necessitating its co-expression and co-purification in the form of the heterodimer with Dbf4. Having screened through a large number of deletion constructs Hughes et al. were able to obtain diffracting crystals of Cdc7 lacking 132 and 46 residues from KI2 and KI3, respectively, bound

to a 141-residue fragment of Dbf4 spanning conserved motifs-M and -C. The crystallized construct also lacked the 36 N-terminal Cdc7 residues, which are highly susceptible to proteolysis *in vitro* and predicted to be disordered. The construct retained only partial activity, and the observed defect was mapped to the deletion in KI2 (Hughes et al., 2012).

Cdc7 has the classic bi-lobal architecture characteristic of protein kinases, comprising N- and C- lobes, with the active site buried in a deep cleft between them (Taylor and Kornev, 2011). The N-lobe (Cdc7 residues 41-135) consists of a an antiparallel β-sheet of 5 strands (β1-5), the αC helix, important for kinase activation, and two non-canonical helices at its N-terminus (N1 and 2) (Hughes et al., 2012) (Figure 1.9). The C-lobe of the kinase is made up largely of α helices and contains a number of features conserved across protein kinases including the catalytic loop containing the RD signature (Arg176, Asp177) and the conserved DFG and APE motifs, which flank the activation loop of the kinase (196-198 and 381-383 respectively). Despite extensive deletions in Cdc7 KI2 and KI3, these motifs were well ordered, and the deletions appeared to have little effect on the overall structure of the kinase. The small portion of KI2 that was visible in the crystals contained an α helix, which packs against the αC helix and provides contacts between the N- and C- lobes of Cdc7. The ordered portion of KI3 contributes a β-strand and two α helices to the canonical C-lobe structure (Hughes et al., 2012).

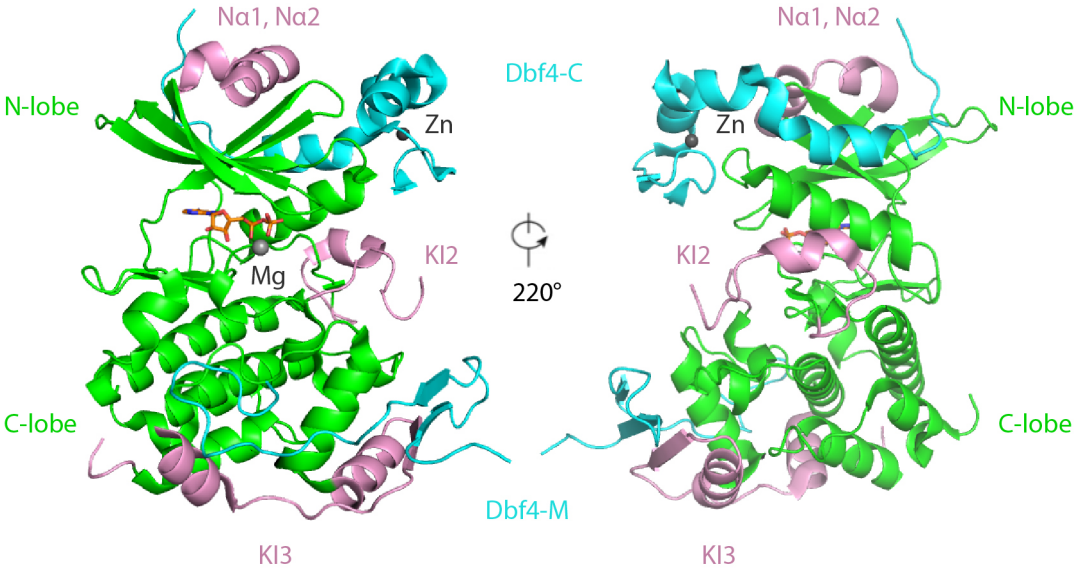


Figure 1-9 Crystal structure of Cdc7-Dbf4 previously determined in the lab

The structure is shown in two orthogonal orientations, with protein chains shown as cartoons. Conserved regions of Cdc7 are shown in green and the non-canonical N-terminus, KI2 and KI3 in pink, and Dbf4 in cyan. Nucleotide (ADP) bound to the active is shown as sticks with carbon atoms in orange with the remaining atoms in standard coloration (blue for nitrogen, red for oxygen and orange for phosphorus atoms); grey spheres are Mg and Zn atoms (Hughes et al. 2012).

The structure revealed that Dbf4 makes a bipartite interaction with Cdc7 with motif-M and motif-C latching onto the C-lobe and N-lobe, respectively. Each interface buried approximately 3,000 $Å^2$ of predominantly hydrophobic molecular surface. Motif-M contains a pair of β strands, of which one forms anti-parallel β-sheet with a β strand from KI3. Motif-M also contains an ordered coiled region that makes extensive hydrophobic interactions with C-lobe. The interactions between KI3 and motif-M were characterised earlier using mutagenesis and site-specific crosslinking (Kitamura et al., 2011).

Motif-C contains a conserved Zn-binding domain, comprised of a single Zn atom coordinated by the side chains of Cys296, Cys299, His309 and His315, which are invariant amongst Dbf4 orthologs. Further to this, motif-C contains three α-helices and two β strands, of which the β strands and the first 2 α-helices are involved in formation of the metal binding site. The Zn-binding domain makes numerous interactions with the N-lobe of Cdc7, of which interactions with αC are particularly important for kinase activation (Figure 1-10) (Hughes et al., 2012). The linker between the two motifs is poorly conserved as with other species. The only

conserved feature is an array of charged residues (288-293) immediately preceding the motif-C. Mutations of these residues were shown to be detrimental to kinase activity, despite them not being ordered in the crystal structure.

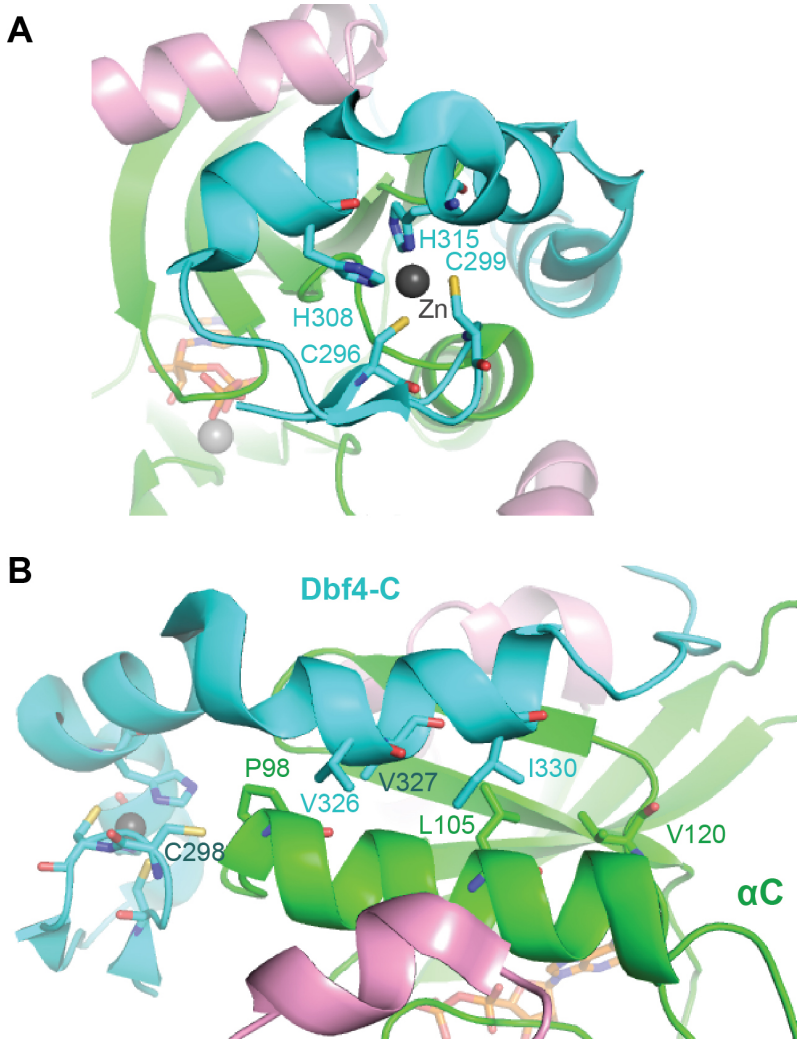


Figure 1-10 Structure of the Dbf4 motif-C Zn-binding site and its interaction with αC

A) Close up view of the Zn-binding domain of Dbf4. The residues responsible for coordination of a Zn atom are indicated. B) Close up view of the interaction between Dbf4 motif-C and the αC helix of Cdc7. Residues involved in the interactions are shown as sticks and labelled.

The activation of Cdc7 was proposed to be analogous to that of the CDKs, where cyclin binding leads to repositioning of the αC helix. This in turn allows for the formation of an essential salt bridge involving a highly conserved Glu residue located within αC and a Lys residue from the kinase AxK motif (Jeffrey et al., 1995). In the case of Cdc7, the salt bridge is formed between Glu104 and Lys90 and this interaction is integral to the structure of the active site. Through producing a range of deletion constructs of the Dbf4 fragment it was shown that motif-C and the preceding charged residues were essential and minimally sufficient for Cdc7 activation. The interaction between the Dbf4 Zn-binding domain and Cdc7 αC were shown to be particularly vital, and point mutations in residues involved in this interaction (V327E and C298E) substantially reduced kinase activity without preventing complex formation. Similar results were also achieved through disruption of the Zn-binding domain itself. The relative protection of a buried cysteine residue in Cdc7 in the presence and absence of motif-C also confirmed that the binding of motif-C stabilises αC of Cdc7 in the active conformation, much like the conformational changes brought about by Cyclin A binding to CDK2 (Jeffrey et al., 1995). Regulation via αC is so common among protein kinases that the helix is sometimes referred to as the Signal Integration Motif (Jeffrey et al., 1995, Yang et al., 2002, Bayliss et al., 2003, Sessa et al., 2005, Taylor and Kornev, 2011). The structure is consistent with a mechanism whereby motif-M tethers Dbf4 to Cdc7 and plays a role in stabilising the C-lobe, and allowing motif-C binding/dissociation to act as an on/off switch of Cdc7 activity. Another interesting observation in the structure was the lack of an activation threonine in Cdc7. In many kinases, including CDK2, there is a threonine residue present in the activation loop, which upon phosphorylation by another kinase, forms part of a hydrogen bond network that is essential for forming the active kinase conformation (Madhusudan et al., 1994, Russo et al., 1996, Canagarajah et al., 1997). In the structure of Cdc7, this threonine appears to be replaced by Glu217 and the previously suggested candidate of Thr376 (Masai et al., 2000) does not overlay with the activation threonine of other kinases.

The structure of Cdc7 also revealed some previously unexpected shared features with other kinases including the MAP kinases, which contain an insert sequence similar to KI3 that was shown to be important for interaction with other proteins

(Chou et al., 2003). MAP kinases also contain a C-terminal extension harbouring a small helix (αL16), which packs against the N-lobe in a similar way to motif-C of Dbf4. Similarly, hydrophobic motifs found within the C-terminal extensions of protein kinases A, B, C, G are involved in the alignment of the αC (Yang et al., 2002, Kannan et al., 2008).

Crystallisation of Cdc7-Dbf4 with two inhibitors revealed the determinants for their binding and explained the increased affinity and selectivity of XL413 over PHA767491 (Figure 1-11). Both inhibitors interacted with the side chains of the conserved Lys90 and Asp196 residues of Cdc7, PHA767491 also interacted with the main chain amide of Leu137, while XL413 engaged the side chain amide of Asn182. All further interactions were accounted for by van der Waals interactions with the nucleotide-binding pocket. XL413 made more extensive contacts with the P-loop of Cdc7, which lead to an inward motion of the loop and closing of the pocket. Both inhibitors also made contacts with residues in the active site that are not conserved among kinases (Met118 and His139) and XL413 made one further contact with Ser70. These increased interactions relative to PHA767491 likely explain the higher potency and selectivity of XL413 (Hughes et al., 2012).

While the crystal structures determined by Hughes et al revealed a great deal of previously unknown information about the mechanism of Cdc7 activation by Dbf4, many questions remained about the roles of KI2 and KI3 in the kinase, and how the kinase facilitates interactions with a number of possible substrates. These questions form a large part of the work in this study.

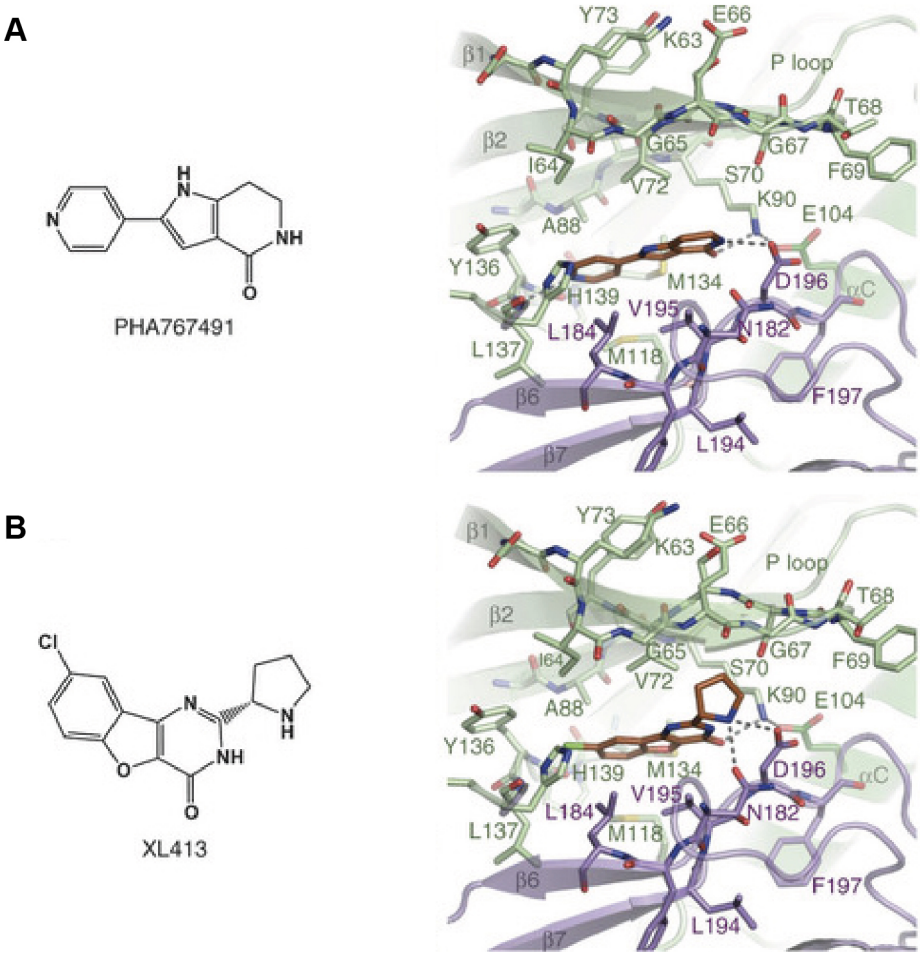


Figure 1-11 Binding of inhibitors to Cdc7-Dbf4

A) Chemical structure of PHA767491 and its binding to the active site of Cdc7. B) XL413. Protein chains are shown as cartoons, with the N-lobe of Cdc7 in green and the C-lobe in purple. Interacting residues are shown as sticks. Inhibitors are shown as sticks in orange with all non-carbon atoms shown with standard colouration. Interactions between amino acid side chains and the inhibitors are shown as dashes (adapted from Hughes et al. 2012).

1.7 Project aims

Although much is already known about Cdc7-Dbf4 and its structure and function, a number of important features for its activity and physiological functions still remain a mystery. With Cdc7 inhibition appearing to be a promising target for drug development, it is important to elucidate the structural and functional features unique to this kinase. The aim of this project was to use X-ray crystallography in tandem with biochemistry and cell biology to structurally and functionally characterise a Cdc7-Dbf4 construct with kinase activity equivalent to that of the wild type (WT) heterodimer. In chapter 3, the design and purification of Cdc7-Dbf4 constructs for use in crystallography is described. In chapters 4 and 5, three structures of an active Cdc7-Dbf4 construct are presented bound to XL413, an ATP mimic and an Mcm2-derived substrate peptide. In chapter 6, the functional relevance of the structural features of Cdc7 revealed in the previous chapters is investigated *in cellula* using an inducible knockout cell line created as part of this study.

Chapter 2. Materials & Methods

2.1 General methods for DNA plasmid construction

2.1.1 Polymerase chain reaction (PCR) for amplification of DNA

PfuUltra II fusion hot start DNA polymerase (Agilent) was used for PCR amplification of DNA templates unless otherwise stated. A typical reaction mixture of 100 μl contained 50-250 ng of template DNA, 250 μM of each deoxynucleotide triphosphate (dNTP) [deoxyribo-adenosine triphosphate (dATP), deoxyribo-thymidine triphosphate (dTTP), deoxyribo-cytidine triphosphate (dCTP) and deoxyribo-guanosine triphosphate (dGTP) with 100 pmol of each oligonucleotide primer and 1μl of enzyme in PfuUltra II reaction buffer supplied with the enzyme. All PCR reactions were carried out using a Thermocycler C1000 instrument (BioRad). A typical PCR programme consisted of one cycle of denaturation at 95°C for 5 min followed by 30 cycles of amplification consisting of a further 30 s of denaturation at 95°C, 30 s primer annealing at 54°C and 2 min/kb template extension at 72°C, ending with a 7-min incubation at 72°C to complete replication of the DNA template.

PCR products were separated by gel electrophoresis in 1-1.5% (w/v) agarose gels made up in TAE buffer (40 mM Tris Base, 20 mM acetic acid, 50 mM ethylenediaminetetraacetic acid (EDTA)) and containing a 1 in 10,000 dilution of GelRed (Biotium) for visualisation. Higher percentage agarose gels (2%) were used for separation of smaller PCR fragments.

2.1.2 Overlap-extension (splice) PCR

This method was used to splice DNA fragments for production of deletion mutants of Cdc7 as well as for the N-terminal fusion of mCherry to Cdc7 for cell line production. Typically, splicing PCR reactions of 100 μl contained 10 μl of each fragment to be fused together, 250 μM of each deoxynucleotide triphosphate (dNTP) and100 pmol of each end primer for the final construct. Reactions were performed in 1xPfuUltra II reaction buffer supplied with the enzyme. 1 μl of the

polymerase was added last before initiating the PCR. The PCR reaction was then carried out as previously described before excision and purification.

2.1.3 DNA purification from agarose gels

Fragments were isolated and purified using a QIAquick gel extraction kit (Qiagen) as per the manufacturers instructions. The excised gel slice containing the desired DNA fragment is dissolved in 3 volumes of QG buffer (supplied) by incubation at 50°C. The high salt conditions force the DNA to bind to the silica-gel membrane filter of the QIAquick spin column, while impurities such as dNTPs, PCR buffer components and protein, pass through and are discarded. The spin column is then washed and the DNA is eluted in distilled water or the low-salt EB buffer (supplied).

2.1.4 DNA digestion for plasmid construction

Five to ten μg of plasmid of PCR product was mixed with 40-80 U (2-4 μl) of the relevant restriction enzyme (New England BioLabs) in the presence of the reaction buffer supplied with the enzyme. Double digests using two enzymes were performed when the reaction buffers were compatible. Reactions were made up to a final volume of 100 μl with distilled water and incubated at 37°C for 2-4 hours. When different reaction buffers were required for the two restriction enzymes, the single digested DNA was purified using a QIAquick PCR purification kit (Qiagen) as per the instructions and used to set up a digestion reaction with the second enzyme. Prior to use, DNA fragments were separated in 1-2% (w/v) agarose and isolated using the QIAquick gel extraction kit (Qiagen) as described previously (2.1.3).

2.1.5 DNA ligation

A 1:1 molar ratio of fragments were mixed and ligated using 1U (1 μl) T4 DNA ligase (Invitrogen) in the reaction buffer supplied with the enzyme (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 1 mM adenosine triphosphate (ATP) and 1 mM dithiothreitol (DTT) in a final reaction volume of 10 μl. Reactions were incubated for at least 2 h, but typically overnight, at 15°C.

2.1.6 Transformation of chemically competent *E.coli* cells

One to five µl of the DNA ligation reaction or 1-100 ng of plasmid DNA was added to 25-50 µl chemically competent *E. coli* cells and incubated on ice for 15 min. The bacteria were then heat shocked at 42°C for 50 sec before cooling on ice for 1 min; 250 µl of SOC medium (casein enzymatic hydrolysate 20 g/L, yeast extract 5 g/L, NaCl 0.5 g/L, $MgSO_4$ 2.4 g/L, KCl 0.186 g/L and 0.4% glucose) was added and bacteria were allowed to recover at 37°C for 30 min with gentle shaking. The culture was then seeded on selective LB agar plates containing the relevant antibiotic. The plates were incubated overnight at 28°C or 37°C and single colonies were used to inoculate liquid LB or TB broth cultures containing antibiotic(s) required for selection.

2.1.7 Colony-PCR

Bacterial colonies were screened directly from selection plates by colony-PCR to identify plasmids with the correctly ligated insert. A pair of primers was selected such that PCR product would span a ligation point. A fraction of each colony was added to a 20 µl PCR mixture containing 125 µM of each dNTP, 10 pmol of each primer, 0.5 units of Taq DNA polymerase (Invitrogen) and 2 mM $MgCl_2$ made up in the reaction buffer provided with the enzyme. As a positive control, 0.5 µl of ligation reaction was used. PCR cycling conditions consisted of 8 min at 95°C to lyse the cells and denature the DNA, 25 cycles consisting of 30 sec of denaturation at 95°C, primer annealing for 30 sec at 54°C followed by 0.5-1 min template extension at 72°C. Reaction products, separated by electrophoresis in 1.5-2% (w/v) agarose, were detected by staining with Ethidium bromide or GelRed. Bacterial colonies containing the plasmid construct of the expected structure produced a DNA band migrating at the expected position, as in the positive control. This method allowed screening large numbers of bacterial colonies in a short time.

2.1.8 Small-scale preparation of plasmid DNA (miniprep)

Plasmid DNA was purified from 3-ml cultures described previously (8.1.5) using commercial QIAprep Spin Miniprep protocol (Qiagen). Bacteria were lysed in buffer P2 (supplied) and the lysate neutralised by addition of buffer N3 (supplied). Buffer N3 also provides the high salt conditions required to precipitate contaminants (such as cellular debris and chromosomal DNA) while maintaining circular plasmid DNA in solution. The lysate, pre-cleared by centrifugation at 10,000 *g* for 20 min, was applied onto a silica membrane with the QIAprep spin column, which binds DNA under high salt conditions. After washing with ethanol-containing PE buffer (supplied), pure plasmid DNA was eluted in 30-50 μl of distilled water or low-salt EB buffer (supplied).

2.1.9 Large-scale preparation of plasmid DNA (maxiprep)

To prepare highly concentrated and pure plasmid DNA for use in transfections, a Plasmid Maxi kit (Qiagen) was used as per the manufacturers instructions. A single bacterial colony was used to inoculate 200-300 ml of LB culture supplemented with the relevant antibiotic(s), which was grown overnight at 37°C with vigorous shaking (250-300 rpm). Bacteria, pelleted by centrifugation and re-suspended in buffer P1 buffer containing RNAse (supplied), were lysed by addition of 10 ml P2 buffer (supplied). The lysate was then neutralised with 10 ml of chilled N3 buffer and precipitated contaminants were removed by centrifugation at 20,000 *g* for 30 min. The clear supernatant, containing plasmid DNA, was loaded onto Plasmid Maxi column pre-equilibrated with QBT buffer (supplied). Plasmid DNA binds to the anion exchange resin under low ionic strength and low pH conditions while impurities are able to flow through. The column was washed twice with 30ml QC buffer (supplied) before elution with 15ml of high salt QF buffer (supplied) and precipitation with 11.5 ml isopropanol. Purified plasmid DNA was then collected by centrifugation and washed with 70% ethanol before dissolving in an appropriate volume of distilled water (0.25-0.5 ml).

2.2 DNA constructs

Plasmids used in this study are listed in Table 8-1 and 8-2. Primers used in PCR and plasmid construction are listed in Table 8-4 (Appendix). Plasmids used in this work but not described in this chapter have been created by others and are referenced in tables 8-1 and 8-2. New England Biolabs supplied all restriction enzymes, and all generated plasmids were verified by sequencing to check for unwanted mutations (LRI sequencing Department). All constructs used for expression in mammalian cells were purified by maxiprep as described previously (2.1.9).

2.2.1 Constructs for bacterial expression of Cdc7-Dbf4 and Cdc7-Drf1 heterodimers

A number of deletion constructs were created of Cdc7 for co-expression with Dbf4 and Drf1 in *E. coli.* The deletions in Cdc7 KI2 and KI3 were created by overlap-extension PCR using constructs previously made by Hughes et al. (2012) as a starting material. The constructs and PCR primers were designed using Vector NTI software (Invitrogen). Reverse and forward deletion primers incorporated 20-30 nt of complementary sequences in their 5' regions to allow PCR splicing. Primer SH78 and the reverse primer for the deletion were used for PCR as previously described. SH78 was used to maintain the ΔN (1-36) deletion, which was maintained in all bacterial expression constructs. The forward deletion primer was then used with primer SH81 for PCR. SH81 is the reverse primer annealing at the very end of the Cdc7 coding region. The fragments produced were separated by agarose gel electrophoresis and gel extracted as previously described. These DNA fragments were then used for overlap-extension PCR and the production of a single DNA fragment with a desired deletion. The spliced fragment, digested with *Xho*I and *Nco*I and purified by gel extraction, was ligated between the *Nco*I and *Xho*I sites of the pRSFDuet1 bacterial expression vector (Novagen) (Figure 2-1). The *Nco*I site is located upstream of the hexahistidine tag (His_6) coding sequence, allowing for expression of untagged proteins.

The Dbf4 MC2 fragment was cloned into a modified version of the pCDFDuet1 bacterial expression vector encoding a recognition sequence for the human rhinovirus 14 (HRV14) 3C protease downstream of the His_6 tag (Hughes et al., 2012) (Figure 2-1). Primers SD25 and SD20 were used as forward and reverse primers, respectively. The resulting fragment was digested with *Xma*I and *Xho*I, purified using agarose gel electrophoresis and ligated between *Xma*I and *Xho*I sites of the modified pCDFDuet1-3C vector. The use of the *Xma*I site allows in-frame fusion of the vector-derived sequence encoding the His_6 tag and the HRV14 3C site with the ligated fragment. Constructs for expression of Drf1 fragments were constructed in a similar way utilising the relevant primers incorporating *Xma*I and *Xho*I sites. For a list of primers used in plasmid production see Table 8-4. Accuracy of all constructs was confirmed by sequencing of small scale DNA preparations (2.1.8) after transformation in *E. coli*.

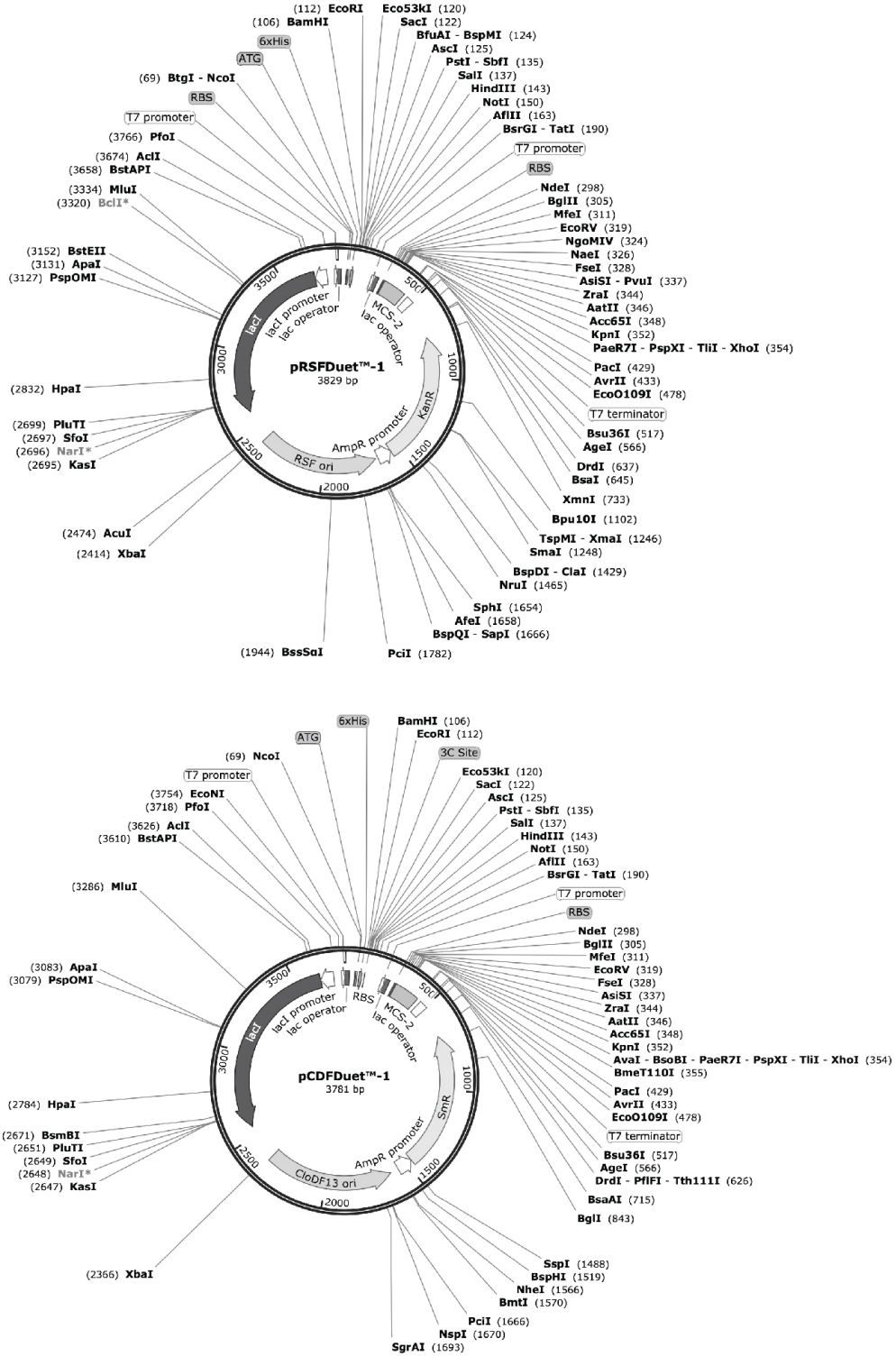


Figure 2-1 Plasmids used for expression of Cdc7 and Dbf4/Drf1 constructs for expression in bacteria

2.2.1.1 Site directed mutagenesis

Site directed mutagenesis was used to create point mutations in Cdc7 for use in *in vitro* kinase assays and in cell line production. Mutations were made in the pRSF-Cdc7 (ΔN) and pBABE-FLAG-Cdc7 vectors. Vectors were mutagenized using the QuikChange procedure (Braman et al., 1996). Primer pairs relevant to each mutation can be found in Table 8-4. For site directed mutagenesis, 100 ng template DNA was denatured at 95°C for 3 min and was amplified for 18 cycles. Each cycle consisted of a 30 s denaturation step at 95°C, primer annealing for 1 min at 54°C and template extension for 12 min at 72°C followed by incubation at 72°C for 15 min. The products, treated with *Dpn*I for 3 h at 37°C to digest the template, were used for transformation of *E. coli*. Mutations were confirmed by sequencing of small-scale plasmid preparations (2.1.8).

2.2.2 Lentiviral and retroviral constructs for expression of Cdc7 in human cells

A number of constructs and mutants of Cdc7 were produced for expression in mammalian cells via retroviral and lentiviral transduction. Plasmids used for mammalian expression can be found in Table 8-2.

A construct for expression of mCherry-Cdc7 fusion was created on the basis of the lentiviral pWPT-GFP vector, a gift from Didier Trono (Addgene plasmid # 12255) (Figure 2-2), which incorporates LoxP sites allowing for excision of the expression cassette by the action of Cre recombinase. Overlap-extension PCR was used to splice a DNA fragment encoding mCherry (produced by PCR with primers SD28 and SD16) and a synthetic codon-shuffled DNA fragment encoding full-length WT Cdc7 (synthetized by GeneArt and PCR amplified using primers SD17 and SD31). Codon-shuffled Cdc7 coding sequence was used to make the transgene resistant to disruption by CRISPR/Cas9 constructs designed to target endogenous *CDC7* alleles. The spliced DNA fragment, digested with *Mlu*I and *Sal*I and purified by agarose gel electrophoresis, was ligated between *Mlu*I and *Sal*I sites of pWPT, replacing the originally present sequence encoding GFP.

A range of DNA constructs for expression of FLAG-Cdc7 variants in human cells were made inserted and into pBABE (puro) (Figure 2-2) (Morgenstern and Land, 1990) and pWZL (hygro) retroviral vectors . PWZL (hygro) was a gift from Scott Lowe (Addgene plasmid # 18750). A DNA fragment encoding FLAG-tagged Cdc7 was PCR-amplified using pQFLAG-Cdc7 (Hughes et al., 2010) as templates and primers SD145 and SD149. The product, digested with *Sna*BI and *Apa*I and purified by agarose gel electrophoresis, was ligated between *Sna*BI and *Apa*I sites of the pBABE (puro) vector. This resulted in the pBABE-FLAG-Cdc7 (WT) construct, which was then used to produce mutants of Cdc7 by site directed mutagenesis and/or overlap-extension PCR (2.2.1.1 and 2.1.2). DNA fragments encoding Cdc7 with Δ2aq mutations were also cloned into the pWZL (hygro) vector by amplification by PCR with SD150 and SD96 and subsequent ligation using the *Sna*BI and *Sal*I sites. For a list of all primers used see Table 8-2.

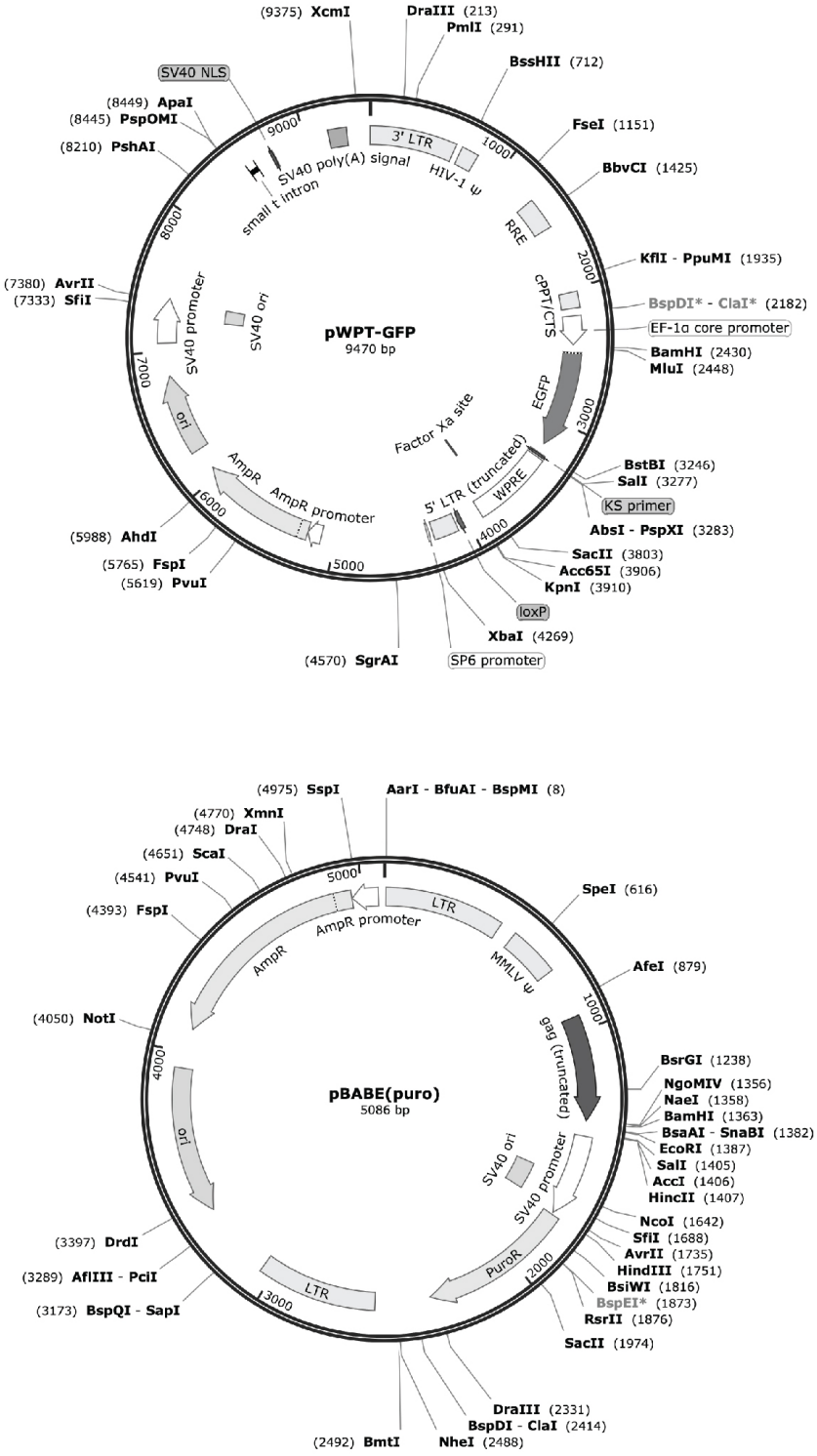


Figure 2-2 Plasmids used for lentivirus and retrovirus production for expression of Cdc7 constructs in mammalian cells

2.3 Cell lines

Cell lines created and used as part of this study can be found in Table 8-3. All cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen) supplemented with 10% foetal bovine serum (FBS, Invitrogen) and 100 U/ml penicillin and streptomycin (Invitrogen). Some cell lines were further supplemented with puromycin or hygromycin as described. Cells were cultured in a humidified, 5% CO_2 atmosphere. HEK-293T cells and HT1080 cells were provided by the LRI Cell Culture Facility (Clare Hall).

2.3.1 Transduction of human cells with lenti- and retroviral vector particles

All cell lines were generated by transduction of cells with lentiviral or retroviral vectors, which were produced by transfection of HEK-293T cells with the designed transfer vector and the relevant packaging (HIV-derived lentiviral p8.2 or MLV-derived retroviral pCG-GAGPOL) and envelope (pMDG, encoding vesicular stomatitis virus glycoprotein G) constructs.

Transfection of HEK-293T cells was performed using X-tremeGENE 9 lipid DNA transfection reagent (Roche) as per the manufacturer's instructions. Cells were seeded in 10-cm culture dishes ($2.8x10^6$ cells per dish) 16 h prior to transfection. Fifteen µl of lipid reagent, diluted in 500 µl Opti-MEM (Invitrogen), was mixed with 5 µg DNA mixture containing transfer, packaging and envelope constructs at a weight ratio of 10:8:2, and lipid-DNA complexes were allowed to form at room temperature for 15 min. During this time, culture media was removed from the HEK-293T cells and replaced with pre-warmed DMEM supplemented with 10% FBS but lacking antibiotics. After 15-min incubation, the 500-µl DNA:lipid mixture was added to the cells in a drop-wise fashion, and the dish was returned to a tissue culture incubator. After 24 h, the culture medium was replaced with fresh medium containing penicillin and streptomycin as previously described. Conditioned medium containing viral vector particles was harvested 48 and 72 h post-transfection and filtered through 45-µm filters to produce a cell-free supernatant containing the desired virus without

any cell debris. The supernatant was then diluted 1:2 with fresh medium and used for infection of HT1080 cells.

Lentiviral vector particles used to make the HT1080 mCherry-Cdc7(LoxP) cell line were produced using the pWPT-mCherry-Cdc7 lentiviral transfer plasmid in conjunction with the lentiviral packaging vector p8.2 and the viral envelope expressing construct pMDG. Two rounds of infection of HT1080 cells were performed 24 h apart to ensure a high efficiency of infection. Cells were then cloned by limited dilution to produce single colonies in 96-well plates, which were expanded for detection of Cdc7 expression by Western blotting (2.4.3). The result cell line was named HT1080 mCherry-Cdc7(LoxP).

Retroviral particles used for expression of FLAG-Cdc7 and its mutant forms in the conditional knockout cell line *CDC7(-/-)* mCherry-Cdc7(LoxP) were produced by co-transfection of HEK-293T cells with WT or mutant pBabe-FLAG-Cdc7 supplemented with the retroviral packaging vector pCG-GAGPOL (Ulm et al., 2007) and the envelope construct pMDG. Conditional knockout cells were subjected to a single round of infection to reduce the chances of overexpression. Cells were treated with puromycin (1 µg/ml). To avoid clonal effects, selected cells were used as mixed populations.

2.3.2 Disruption of endogenous CDC7 alleles using the CRISPR/Cas9 system

Four gRNAs targeting exons one or two of the *CDC7* gene were cloned into the pX459 vector (Ran et al., 2013). To increase the odds of obtaining a frame shift with a complete knock out of the kinase expression, the gRNA target sites were selected immediately downstream of the Cdc7 start codon. HT1080 mCherry-Cdc7(LoxP) cells were transfected with the CRISPR/Cas9 constructs as previously described (2.3.1) and grown for 48 h before addition of 1 µg/ml puromycin. The antibiotic selected for successfully transfected cells, increasing the frequency of gene modification in the surviving cell population. After 24 h in the presence of puromycin, cells were diluted for single colonies and after expansion were tested

for Cdc7 expression by Western blotting. Any apparent knockouts were verified by amplification of the part of the gene targeted by the guide RNA by PCR, which was subsequently cloned into a Topo-TA sequencing vector (Life technologies). This was then used to transform NEB-5α competent *E.coli.* Bacteria were the plated out for single colonies on selective agar plates and small-scale plasmid preparations were made for a number of colonies for each clone. These were sequenced at the in house sequencing facility (London Research Institute) to verify disruption of the *CDC7* gene. The resultant cell line after *CDC7* knockout was named HT1080 (*CDC7*-/-)mCherry-Cdc7(LoxP).

2.3.3 Transgene excision from the conditional knockout cell line for cell cycle assays

As part of the assays used to test Cdc7 mutants *in cellula,* the mCherry-Cdc7 cassette, flanked in the pWPT backbone by a pair of LoxP sites, can be removed by expression of Cre recombinase. To achieve this, cells were harvested and counted before seeding in 12-well dishes (70,000 cells/well). Adenoviral vector for expression of Cre recombinase and GFP (Ad-Cre-GFP) or a control vector expressing GFP alone (Ad-GFP) was applied to the cells at a concentration of 5 x10^6 PFU/ml in 0.5 ml of cell growth medium. The cell-virus mixture was transferred to the tissue culture incubator and agitated every 20 min for the initial 2 h of incubation. Twenty-four h post infection the media containing the adenovirus was removed and replaced with fresh media; the cells were then expanded for downstream analyses as needed.

2.4 Protein analyses

2.4.1 Sodium dodecyl sulphate – polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE gels used for separation of protein samples were composed of a 4% stacking gel layered on top of a 11% running gel. Resolving gels were prepared by supplementing the monomer mix of 11% acrylamide, 0.29% bisacrylamide (w/v), 0.1% sodium dodecyl sulphate (SDS) and 0.38 M Tris-HCl pH 8.8) with 0.042%

(w/v) ammonium persulphate (APS) and 0.84% (v/v) N,N,N',N'-Tetramethylethylenediamine (TEMED). The stacking gel mix was prepared by supplementing a monomer mix of 4% acrylamide, 0.1% bisacrylamide, 0.1% (w/v) of SDS and 0.12 M Tris-HCl pH 6.8 with 0.046% (w/v) APS and 0.123% (v/v) TEMED. Commercial gradient 4-20% Tris-glycine gels (BioRad) were used in some instances.

Protein solubilised in Lamelli buffer (2% (w/v) SDS, 10% (v/v) glycerol, 0.0025% bromophenol blue, 62.5 mM Tris-HCl pH 6.8) supplemented with 25 mM β-mercaptoethanol (BME) were separated by SDS-PAGE in 11% or 4-20% Tris-Glycine gels run at 200 V in Tris-Glycine-SDS (TGS) buffer (25 mM Tris base, 250 mM glycine, 0.1% (w/v) SDS).

2.4.2 Coomassie staining

Proteins separated by SDS-Page were visualised by staining with Coomassie. Gels were rocked in Coomassie staining solution (40% (v/v) methanol, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie Blue R250) until protein bands were visible. Staining solution was then replaced with de-staining solution (40% (v/v) methanol, 5% (v/v) acetic acid) to achieve a clear background.

2.4.3 Western blotting

2.4.3.1 Preparation of whole cell extracts

Cells harvested by trypsinisation were washed with phosphate buffered saline (PBS) supplemented with 0.1 mM phenylmethylsulphonyl fluoride (PMSF) and lysed in modified RIPA buffer containing 0.1% SDS, 0.5% sodium deoxycholate, 1% nonidet P40 (NP40), 150 mM NaCl, 50 mM Tris-HCl pH 8.0, 0.1 mM PMSF and EDTA-free protease inhibitor cocktail (Roche). When blotting with phospho-specific antibodies the RIPA solution was supplemented with phosphatase inhibitors (1 mM β-glycerophosphate and 1 mM NaF). Insoluble debris were removed by centrifugation at 13,000 *g* for 5 min and analysed by SDS-PAGE and Western blotting. Protein concentration of lysates was measured by bicinchoninic acid

(BCA) protein assay (Pierce) with a BSA standard as per the manufacturers instructions.

2.4.3.2 Electrotransfer and detection

Proteins separated by SDS-PAGE were electroblotted onto Trans-Blot Turbo Mini PVDF membranes (BioRad) using a Transblot Turbo semi-dry transfer unit (BioRad) for 30 min at 25 V. Membranes were blocked at room temperature in blocking solution containing 5% (w/v) skimmed milk power in PBS supplemented with 0.1% Tween-20 for at least 2 h to reduce non-specific antibody binding. The membranes were then incubated with appropriate primary antibodies diluted in blocking solution for 2 h at room temperature with gentle rocking. The membranes were then washed in three changes of PBS supplemented with 0.1% Tween-20 for a total of 45-60 min to remove any excess antibody before incubation with the appropriate horseradish peroxidase (HRP) conjugated secondary reagent. Milk powder and phosphate buffers were avoided when using phosphospecific antibodies. The membranes were blocked and incubated with antibodies diluted in 3% (w/v) BSA (Melford) in TBS/Tween (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween-20). Membranes extensively washed to remove unbound secondary reagent were soaked in ECL-prime or ECL-Select chemiluminescent detection reagents (GE Healthcare) prepared as per the manufacturer's instructions. Amersham Hyperfilm ECL films exposed to membranes were developed using automatic film processing system (JP-33, JPI).

2.4.3.3 Stripping of PVDF membranes

Membranes were rinsed in PBS/Tween to remove the detection reagent before incubating in stripping buffer (71 mM β-Mercaptoethanol, 2% (w/v) SDS, 62.5 mM Tris-HCl pH 6.8) for 30 min at 50°C. After several rinses with PBS/Tween the membrane was blocked and re-probed as previously described (2.4.3.2).

2.4.3.4 Antibodies used for Western blotting

The following primary antibody dilutions were used for Western blotting: Mouse anti-Cdc7 (Neomarkers) 1:10,000, rabbit anti Mcm2-phospho-S40 (Abcam) 1:1,000, goat anti-Mcm2 (N19) (SantaCruz) 1:500 and mouse anti-β-tubulin (Abcam) 1:5,000.

Horseradish peroxidase-conjugated anti-rabbit, anti-mouse and anti-goat secondary antibodies (Abcam) were all used at a dilution of 1:10,000.

2.4.4 Production and purification of Cdc7-Dbf4 and Cdc7-Drf1 constructs

All Cdc7 mutants were expressed using the pRSF Duet1 bacterial expression vector (Novagen) without a tag. All Dbf4 and Drf1 constructs were produced in a His_6-tagged form of pCDF-His6-3C (Hughes et al., 2012), which is a modified version of the pCDF Duet1 vector (Novagen). Crucially, the pCDF and pRSF vectors contain compatible origins of replications, allowing for the plasmids encoding Cdc7 and Dbf4 or Drf1 to stably co-exist and doubly transformed bacteria can be selected in the presence of spectinomycin and kanamycin.

E. coli Rosetta-2 (DE3) cells (Novagen) co-transformed with the respective versions of pRSF-Cdc7 and pCDF-His6-3C-Dbf4 or pCDF-His6-3C-Drf1 and were used for expression of the recombinant heterodimers. Bacteria were grown in LB supplemented with 50 μg/ml kanamycin and 100 μg/ml spectinomycin in shake flasks at 30°C. Cells were then cooled to 18°C upon reaching an A_{600} of 0.8, supplemented with 50 μM $ZnCl_2$, and protein expression was induced by addition of 0.01% w/v isopropyl-β-D-thiogalactopyranoside (IPTG). Following incubation overnight (~16 h) at 18°C with vigorous shaking, the cells were harvested by centrifugation and stored at -80°C.

For purification, cells were thawed and re-suspended in kinase core buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10% glycerol, pH 7.5) supplemented with 0.1 mM PMSF, complete EDTA free protease inhibitor mix, 1 mg/ml lysozyme and 0.5% NP40. The cells were disrupted by sonication on ice and the lysate was clarified by

centrifugation at 17,000 *g* at 4°C for 30 min using a Sorvall SS-34 rotor. The clarified extracts, supplemented with 20 mM imidazole were incubated with 3 ml of Ni-NTA agarose (Qiagen) for 30 min at 4°C with gentle rocking. The resin, collected by filtration through an empty gravity column (BioRad), was washed with 4 changes of 30 ml kinase core buffer supplemented with 20 mM imidazole to remove unbound and weakly bound proteins. His_6-tagged proteins were eluted with 10 ml of kinase core buffer supplemented with 200 mM imidazole in 1-ml fractions. Protein containing fractions were pooled, diluted with salt-free buffer to adjust NaCl concentration to 200 mM NaCl, supplemented with His_6-tagged λ-protein phosphatase (1 mg per 15 mg of protein), His_6-tagged HRV14 3C protease (1 mg per 50 mg of protein), 1 mM $MgCl_2$, 2 mM $MnCl_2$ and 1 mM DTT and incubated overnight at 4°C to allow for His_6-tag cleavage from the Dbf4 or Drf1 mutant and de-phosphorylation of the proteins. The next day, precipitated material was removed by centrifugation and the soluble protein was dialysed against a large excess of low salt/imidazole buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 20 mM imidazole, 2 mM DTT) for 2 h at 4°C using a Slide-A-Lyzer cassette (Pierce). Dialysed protein was passed through a 1-ml HisTrap FF column to remove His_6-tagged phosphatase and protease. The recombinant Cdc7-Dbf4/Drf1, collected in the flow through, was diluted 1:5 in salt-free buffer to adjust NaCl concentration to ~80 mM and filtered through a 1-ml HiTrap Q HP column to remove further contaminants. The heterodimers, collected in the flow through, were supplemented with 2 mM DTT and concentrated to ~5 ml using a VivaSpin 20 centrifugal concentrator with a 10,000 MWCO and PES membrane (Viva products). Complexes were then further purified by size exclusion chromatography on a Superdex 200 16/600 column operated in 150 mM NaCl, 25 mM Tris-HCl, pH 7.5. Fractions were analysed by SDS-PAGE and those containing Cdc7-Dbf4/Drf1 heterodimers were pooled and dialyzed against excess of a $ZnCl_2$ containing buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM DTT, 75 μM $ZnCl_2$) overnight at 4°C. The protein was then concentrated to ~10 mg/ml and either used in crystal trials or supplemented with 10% glycerol and flash-frozen in liquid nitrogen for later use in kinase assays.

2.4.5 Crystal screens

Crystal screening was performed using a range of commercially available screens of sparse matrix conditions. Such screens typically contain 96 conditions that have been identified as being previously successful in the crystallisation of other proteins through data mining. Many screens are designed for the crystallisation of specific protein families. Screens were prioritised for screening, which were specifically designed for crystallisation of kinases (JB Kinase) as well as prioritising those, which had contained promising conditions in previous experiments. Screens were utilized in the order presented in the table until the entire prepared protein sample had been utilised.

Table 2-1 Table of commercially available screens used in this study

Screen Name	Manufacturer	Product No.
JB Kinase	Jena Bioscience	CS-204L
JB cryo	Jena Bioscience	Discontinued
JCSG+	Molecular Dimensions	MD1-37
PACT Premier	Molecular Dimensions	MD1-36
Wizard Cryo 1+2	Molecular Dimensions	1008649
Wizard 1+2	Molecular Dimensions	1008647
Wizard 3+4	Molecular Dimensions	1008648
Morpheus	Molecular Dimensions	MD1-47
Peg Ion	Hampton Research	HR2-139
Crystal Screen	Hampton Research	HR2-130
Midas	Molecular Dimensions	MD1-59
Natrix	Hampton Research	HR2-131
MPD Suite	Qiagen	130706

2.4.6 Synthesis of ATPγS conjugated peptides

All ATPγS-conjugated peptides used in this study were synthesised in the in-house peptide synthesis facility based at the London research institute (LRI), London. All

Fmoc-protected amino acids, N-α-Fmoc-N-β-4-methyltrityl-L-diaminopropionic acid and the Rink Amide resin were purchased from Merck. Adenosine 5'-O-(3-Thiotriphosphate) tetralithium salt and Adenosine 5'-O-(2-Thiodiphosphate) trilithium salt were purchased from Calbiochem. Dichloromethane (DCM) and Dimethylformamide (DMF) were purchased from VWR. Acetonitrile (ACN) and Diethyl ether were purchased from Fisher. Bromoacetic acid, Triethylammonium bicarbonate buffer, Trifluoroacetic acid (TFA), Triisopropylsilane (TIS), and Diisopropylcarbodiimide (DIC) were purchased from Sigma.

Peptides backbone were synthesised on a 433A peptide synthesiser by solid phase peptide synthesis following standard Fmoc chemistry. Peptides were assembled on a Rink Amide resin (LL 0.30 – 0.40 mmol/g) in order to have a C-terminal amide. ATP-peptide conjugates are usually used as bisubstrate analogue inhibitor of the serine/threonine kinase proteins. In this work serine was replaced with aminoalanine using the derivative N-α-Fmoc-N-β-4-methyltrityl-L-diaminopropionic acid (Fmoc-Dap(Mtt)-OH). After synthesis, the side-chain Mtt group was selectively removed by treatment of the peptidyl resin with 1% TFA, 4% TIS in DCM. Afterwords a bromoacetyl moiety was linked to the aminoalanine by reacting the peptidyl-resin with 8.3 eq of bromoacetic acid, 8.3 eq of DIC in a minimum amount of DMF under stirring condition for 2 hours at RT. Peptides were then cleaved from the resin by treatment with 10 ml (v/v) of 95% TFA, 2.5% TIS, 2.5% water for 2 hr. Peptides were precipitated with diethyl ether, dissolved in water after centrifugation and lyophilized. Purification was performed by reversed-phase HPLC using 1% ACN, 0.08% TFA in H2O (buffer A) and 90% ACN, 0.08% TFA in H2O (buffer B). The desired mass of the peptides was confirmed by LC-MS. The ATPγS/ADPγS-peptide conjugates were obtained through a thioether bond formation by reacting the sulfhydryl group of ATPγS and the bromoacetyl moiety linked to the aminoalanine.

In previous methods (Parang et al., 2001) the purified bromoacetylated peptide was dissolved in a solution MeOH:H2O 4:1 and treated with ATPγS (adenosine 5’-O-3-thiotriphosphate) for 24h at room temperature. The resulting peptide was purified by reversed phase HPLC using a gradient of H2O:ACN under neutral conditions without using TFA. For peptides used in this study, Triethylammonium bicarbonate

buffer (1M, pH 8.4-8.6) was used to perform this reaction since the bromoacetylated peptides were not soluble in the MeOH:H2O 4:1 solution. Peptide and ATPγS were dissolved separately in the buffer in a ratio 1:2 respectively. The two solutions were then mixed and the reaction was carried out for 24 h at room temperature. The compound was lyophilised 2 or 3 times until complete removal of the salt by addition of water. Peptides with pI > 3 were purified by reversed phase HPLC under neutral condition. Peptides with pI < 3 when dissolved in water were showing auto-hydrolysis of the ATPγS group, as this is very acid sensitive. Therefore these peptides were purified using an Extend-C18 column, which can withstand high pH (up to 11.5) compared to standard C18 column. These peptides were purified under a gradient of 10mmol NH4 (buffer A) and 90% ACN-10 mmol NH4 (buffer B). The purity of the final product was verified using LC-MS in negative ion mode. Overall, for most of the peptides, the use of Triethylammonium bicarbonate buffer resulted more convenient and efficient than the MeOH:H2O solution.

2.4.7 *In vitro* kinase assays

Cdc7 kinase assays were carried out using Mcm2-derived peptide substrates synthesized in the in-house peptide synthesis facility (Lincolns Inn Fields, London). The substrate peptides spanned residues 35-47 of human Mcm2 (TDALTS[pS]PGRDLP-biotin) and were biotinylated at the N-termini. Unless stated the peptide contained a phosphate group attached to Ser41 [pS] residue, which is phosphorylated by CDK2 *in vivo* (Montagnoli et al., 2006). Where required by the experiment, substrate peptides were synthetized with mutations as indicated (Table 8-5). This assay was developed by Hughes et al. (2012).

Each 25-µl kinase reaction mixture contained 5 µg biotinylated substrate peptide and 2.8 nM Cdc7-Df4 or Cdc7-Drf1 or corresponding molar equivalents of a mutant form with 3 µCi of [γ-^{32}P]ATP (3,000 Ci/mmol) in 10 mM $MgSO_4$, 2 mM DTT, 1 mM β-glycerophosphate, 1 mM NaF, 80 µg/ml BSA, 0.1% NP-40, 0.1 mM ATP, 40 mM Hepes-NaOH, pH 7.4. Kinase reactions, were incubated for 30 min at 30°C before

stopping with the addition of guanidine hydrochloride to a final concentration of 2.5 M. The reactions were then spotted onto SAM^2 biotin-capture membranes (Promega) before sequential washes with 2 M NaCl solution (three times), 2 M NaCl in PBS (four times) and distilled water (twice). Each washing step was performed with rocking for a minimum of 2 min. The membranes were then washed with 95% ethanol and air-dired. Incorporated radioactivity was detected and quantified by exposing the membranes to a phosphor storage screen which was then scanned using a Storm-860 phosphorimager (GE Healthcare).

2.4.7.1 In vitro kinase assays with immobilized peptide arrays

The membranes containing immobilized peptides derived from human Mcm2, Mcm4 and Mcm6 were briefly soaked in kinase assay buffer (50 mM Tris pH7.5, 10 mM $MgCl_2$, 2 mM DTT, 1 mM β-glycerophosphate, 1 mM NaF and 0.1% NP40) then washed with sequential changes of 50%, 25%, 12.5% and 5% methanol in distilled water followed by a 20 min incubation in distilled water. Following this procedure, all visible peptide spots on the membranes appeared hydrated. The membranes were washed twice with an excess of kinase buffer and processed for the *in vitro* phosphorylation assay.

Each membrane was soaked in 6 ml kinase buffer supplemented with 80 µg/ml BSA, 700 µCi of [γ-^{32}P]ATP. Recombinant Cdc7(ΔN)-Dbf4(MC) (1.4 µg) was then added to the reaction and the membrane was rocked at 30°C for 30 min. The assay was stopped by careful aspiration of the reaction mixture and soaking the membrane with 6 ml of 0.5 M EDTA for 10 minutes. The membrane was the washed sequentially with rocking in 1 M NaCl (20 min), 1% SDS (20 min), 2 M NaCl (twice for 10 min) and distilled water (twice for 10 min). The membrane was then washed 15 times (5 minutes per wash) in 0.5% phosphoric acid before leaving rocking in 20 ml 0.5% phosphoric acid over night at room temperature. The next day the membrane was rinsed with distilled water followed by submersion in 95% ethanol for 15 s and air-dried. Dry membranes were analysed by phosphorimaging using a Storm-860 instrument (GE Healthcare).

2.5 Cell cycle analysis using flow cytometry

Cells were stained for actively replicating DNA using a Click-iT Plus EdU Alexa Fluor 647 Flow Cytometry Assay kit (Invitrogen) and for total DNA content using propidium iodide. Sample preparation was carried out according to the manufacturer's instructions. Cells in culture in 10-cm dishes were incubated with 10 μM EdU (supplied) by addition of a concentrated stock solution to the culture media for 2 h prior. Cells harvested by trypsinization and washed with PBS supplemented with 1% BSA were (PBS/BSA) fixed by re-suspension in 100 μl of Click-iT® fixative solution containing 4% paraformaldehyde in PBS for 15 min before washing again in PBS/BSA. Cells were then permeabilised by re-suspension in 100 μl 1X Click-iT saponin-based reagent (supplied, diluted 1:10 in PBS/BSA). The remaining steps were carried out away from direct light. After 30-min incubation, the cells were supplemented with 500 μl of the Click-iT Plus reaction cocktail (438 μl PBS, 10 μl copper protectant, 2.5 μl picolyl azide fluorescent dye and 50 μl of a 1:10 dilution of reaction buffer additive) to each sample and left for 30 min. In this step the fluorescent dye is covalently attached to the EdU nucleotides incorporated into the newly synthetized DNA. Cells were then washed with 3 ml of Click-iT saponin-based reagent, pelleted and re-suspended in 100 μl of the permeabilisation reagent. Cells were then stained with propidium iodide by addition of 500 μl of permeabilisation solution supplemented with 0.25 μg/ml propidium iodide and 0.1 mg/ml RNAse A (Sigma-Aldrich). Cells were incubated at 37°C for 30 min before analysis by flow cytometry using a FACSCalibur system (BD Biosciences). Alexa Fluor 647 was excited at 635 nm and detected with a 661 nm band pass filter. Propidium Iodide was excited at 488 nm and detected using a 585 nm band pass filter. The data were acquired using CellQuest software (BD Biosciences) and analysed in FlowJo v9.9. Cell populations were gated using the plot of forward scatter versus side scatter. Single cells were then discriminated from doublets by gating on the PI area vs PI height plots. These gated cells were then used for analysis of cell cycle profiles.

2.6 X-ray crystallography methods and theory

X-ray crystallography is a well-established method for studying structures of macromolecules. Consequently numerous comprehensive texts have been published on X-ray crystallographic theory and practice (Rupp and Kantardjieff, 2010). This section contains a brief overview of X-ray crystallographic theory and the techniques and computational processes used in this study.

2.6.1 Production of protein crystals

2.6.1.1 Crystal growth and optimisation

Formation of crystals is achieved through inducing a highly purified protein preparation to enter a state of limited supersaturation that is prone to crystal nucleation and subsequent growth (Figure 2-3). There are 3 possible phases of supersaturation:

1. Nucleation zone – excess protein can spontaneously form into new crystalline structures.
2. Metastable zone – nucleation cannot occur to form new crystals, but currently existing crystals are able to grow.
3. Precipitation zone – excess protein forms amorphous aggregates. These aggregates are of no use in crystallography.

Crystal trials are used to sample supersaturation space and identify conditions, in which a given protein will enter the nucleation zone and subsequently move into the metastable zone, forming crystals without aggregating. The conditions needed to achieve this are specific to each protein and, therefore, thousands of conditions are often tried before gaining a single crystal.

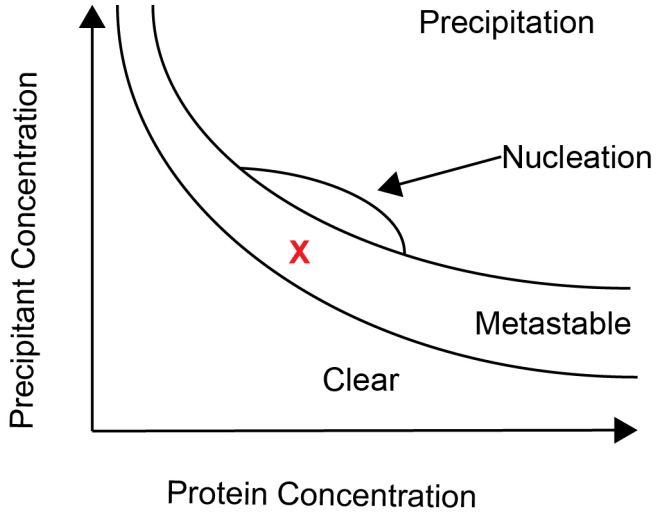


Figure 2-3 Crystallisation by vapour diffusion

X=optimal protein/precipitant concentration for crystal growth.

A common method of sampling supersaturation space is vapour diffusion, whereby a small volume (0.05 -1 μl) of concentrated (5-20 mg/ml) protein solution is added to an equal volume of reservoir solution containing a precipitant. This drop is then sealed in a container with a large excess of reservoir solution and maintained at a constant temperature with minimal vibration. The protein-containing drop will equilibrate with the more concentrated reservoir solution through a vapour phase, which will usually slowly reduce the volume of the drop, causing the protein to sample super saturation space until the solutions have equilibrated. Typically, a wide range of reservoir solutions must be tested to identify optimal crystallisation conditions. A number of variables can be optimised, including pH and type of buffer, precipitant type and concentration, protein concentration and drop volume, temperature and nature of chemical additives. Upon identifying a suitable crystallisation condition, crystals are optimised by creating fine grids of conditions, in which each component of the reservoir solution is varied against the precipitant concentration to produce better diffracting crystals. At this point, it is also common to try adding a number of different chemical additives, which may improve crystal quality. Additive screening was integral to the optimisation of crystals obtained in this study, as was optimisation of the pH of the reservoir solution.

An important crystallography technique often used for production of crystals is microseeding, whereby small crystal fragments are used to nucleate crystal growth in conditions that would normally not allow spontaneous nucleation. In this study, crystals of Cdc7-Dbf4 bound to XL413, which could be obtained by spontaneous nucleation, were transferred to new protein-reservoir drops supplemented with alternative ligands. This process allowed the second series of drops to have a sufficiently low precipitant concentration to equilibrate directly into the metastable zone and produce crystals that were otherwise impossible to obtain.

2.6.1.2 Cryo-protection of crystals

Due to X-rays being an ionising form of radiation, crystals are subject to a large amount of radiation damage during data collection. Radiation damage in crystallography is present in two forms. Dose-dependent radiation damage is postulated to result from the ejection of electrons from the electron spheres of a molecule due to its interaction with X-ray photons (Henderson, 1990). Radiation damage can also be time dependent and this is thought to arise from the accumulation of free radicals within the protein crystals. Free radicals can damage the crystallised molecule and the large solvent content of protein crystals acts as a medium for their dissemination (Matthews, 1968). For this reason, experiments with crystals of macromolecules are almost always carried out at low temperatures (~100 K) to reduce radiation damage by minimising the movement of free radicals through the solvent (Haas and Rossmann, 1970). This drastically reduces the time dependent radiation damage in a given experiment, and even though dose dependent damage remains the same, the reduction in X-ray damage is usually sufficient to allow for a full data set to be collected from a single crystal.

To prepare crystals for data collection, crystals are cryo-cooled by plunging into liquid nitrogen. However, it is crucial to avoid formation of crystalline ice during this process. Crystalline ice not only damages the crystal, but degrades diffraction data due to its ability to diffract X-rays. Therefore, cryoprotectants are used to create conditions that promote formation of amorphous ice. In macromolecular crystallography experiments, cryoprotectant solutions are typically comprised of a

solution similar to the crystallisation condition that has been supplemented with a cryoprotecting agent such as glycerol, polyethelene glycol (PEG) 400, ethylene glycol or 2-methyl-2,4-pentanediol (MPD). The optimal concentrations of a cryoprotectant and soaking time are determined experimentally and vary from crystal to crystal. It is also often necessary to adjust concentration of the cryoprotectant by gradually increasing its concentration to avoid crystal shrinking or cracking, both of which can have a detrimental effect on X-ray data quality Cryocooling also allows samples to be stored for extensive periods of time before data collection.

2.6.2 Data collection and processing

Prior to freezing, crystals are mounted on nylon loops, which create little in the way of background signal from X-ray diffraction. When collecting data, the loops are mounted either manually or by a robot onto a rotating goniometer head and aligned in the path of an X-ray beam, while a stream of liquid nitrogen vapour maintains the temperature of the crystal. The X-ray beam is directed at the rotating crystal and the X-rays are scattered by the electron shells of the macromolecules forming the crystal lattice. The diffraction data recorded on a detector can be subsequently used to calculate an electron density map of the crystallised material.

2.6.2.1 Basic principles of X-ray diffraction

X-rays are primarily scattered through interaction with the electron shells of the atoms of the crystallised molecule. The angles of these diffracted beams can be calculated if diffraction is treated as reflections from equivalent parallel planes designated by three numbers *h, k, l,* where evenly spaced parallel planes in relation to the unit cell are defined by the miller indices *hkl*. Braggs law gives the angles for coherent and incoherent scattering from a crystal lattice, in which X-rays of wavelength λ are reflected by a set of parallel planes (*hkl)* with a spacing of d_{hkl} between planes at a given angle of incidence (θ).

$$2d_{hkl}\sin\theta = n\lambda$$

Equation 2-1 – Bragg's Law

Reflected waves can be in phase and constructively interfere, creating a diffracted beam, which is recorded by a detector. Constructive diffraction occurs when n is an integer at angle θ. The minimum inter-planar distance that creates constructive diffraction defines the maximum resolution of the collected data. When waves are not in phase they cancel each other in a process called destructive interference, which does not create a diffracted beam. The size of the unit cell of the crystal defines how many diffracting planes are present, with unit cells containing more planes producing more reflections.

2.6.2.2 X-ray data collection

In this thesis, all diffraction data was collected at synchrotron sources. A synchrotron uses a particle storage ring as a source of X-rays. Within this ring, powerful magnets accelerate electrons until they approach the speed of light. When accelerated electrons are forced into a curved path, they emit energy in the form of electromagnetic radiation at a number of different wavelengths. Monochromators and a series of mirrors are used to isolate and focus beams of X-rays at a required wavelength. This allows X-rays to be tuned for specific uses in diffraction experiments. Most protein crystal diffraction experiments use X-rays with wavelengths of 0.9-1 Å. Synchrotron sources allow for much faster data collection and detection of weaker reflections than would be possible using laboratory-based instrumentation due to the significantly greater intensity of the X-ray beam and the availability of cutting-edge detectors at the national shared facilities.

A number of methods are available for detection of X-ray diffraction data. In this thesis, data collected at the ESRF beamline BM14 were acquired on a charge coupled device (CCD). A CCD detects visible light, which is created by a phosphor screen placed in front of the detector. This screen creates visible light proportional to the intensity of diffracted X-rays, which can then be read by the CCD. Data

collected at the Diamond Light Source beamline I03 used a new-generation Pilatus 6M detector. This is a pixel device, in which X-rays are converted directly to electric signal by the photoelectric effect in silicon subject to a bias voltage. The electrical signal can then be rapidly detected by an application-specific integrated circuit. This drastically reduces the time required to collect a dataset by reducing the changeover times between images as well as allowing for shorter exposure times due to the increased sensitivity of the system. This, in turn, decreases the amount of radiation damage the crystal is subjected to, allowing for fast collection of high-redundancy data.

2.6.2.3 Mosaicity and oscillation

Crystals are made up of an ordered lattice of molecules. In a perfect crystal, every unit cell would be identical and aligned at the same angle. However, in a typical crystal unit cells can vary slightly in size leading to imperfections in the alignment of the crystal lattice. Such heterogeneity is referred to as mosaicity, which is a prominent problem in macromolecular crystallography, as the crystallised molecules are relatively flexible and held together by weak interactions. Mosaicity leads to reflections that are spread as the crystal is rotated and can lead to partial reflections, in which the reflection is split over a number of frames. The oscillation angle between individual frames is an important factor to consider during data collection, as it is important to minimise the number of overlapping reflections. If the unit cell of a crystal is very large, reducing the oscillation to a very small angle can be helpful in reducing the number of overlaps. The choice of oscillation angle is also influenced by the X-ray source and detector used for the experiment. A small oscillation angle requires more images for a complete data set, which can be prohibitive with weaker sources, and slower detectors. A larger oscillation angle may help to increase the signal-to-noise ratio. Consequently, oscillation angles are determined for each individual experiment to acquire an optimal data set.

2.6.2.4 Selecting oscillation range for data collection

The number of individual frames that need to be collected is crystal-dependent, with those containing a higher order symmetry requiring less rotation/frames to achieve completeness (*i.e*. acquisition of ~100% possible diffraction spots for a given resolution range). The required oscillation range required can be predicted from a small number of test exposures (2-3 diffraction images separated by 90^{o}), which can be processed using Mosflm (Leslie and Powell, 2007) or other indexing software. Mosflm determines the unit cell dimensions and predicts symmetry directly from the diffraction pattern. This process was straightforward for all the crystals obtained in this study, which had tetragonal symmetry ($P4_12_12$) and similar unit cell dimensions. Due to the current speed of data collection it is common to collect more data than absolutely needed. This not only increases the signal to noise ratio of the data collected, but supplies more data for processing should the space group have been incorrectly calculated at a higher order of symmetry.

2.6.2.5 Data processing

Data processing for all structures was completed using the Xia2 automated pipeline (Winter, 2010a). Xia2 utilises the programmes XDS (Kabsch, 2010a) for indexing and integration, Pointless and Scala (Evans, 2006a) for determination of symmetry and scaling. XDS refines the size and orientation of the unit cell and determines possible space groups for the crystal. Pointless uses intensities of integrated reflections to determine point group and make a reasonable guess of a space group, based on systematic absences. Modern software is usually very good at determining the space group, however, errors in this process can sometimes be revealed at subsequent steps that can require the data to be re-processed. In less trivial cases, the initial space group assignment may need to be re-evaluated at a later stage, even during model refinement.

Following indexing, peak intensities are integrated by XDS and the reflections can then be corrected for a number of experimental errors. Factors that can affect integration include crystal decay from radiation damage, incident beam polarisation and relative differences in sensitivity on different parts of the detector. Scala also

calculates standard deviations for equivalent reflections. The data quality in this respect is described by the R_{merge} value, which describes the agreement of symmetry-related reflections with respect to the resolution of the data. R_{merge} is used in conjunction with the $<I/\sigma(I)>$ parameter (average signal-to-noise ratio) to decide the maximum resolution of data, which is of sufficient quality for subsequent processing. R_{merge} is inversely related to the redundancy of the data, meaning that data quality appears to decrease as each reflection is measured multiple times. However, the accuracy of the mean intensity value is increasing, meaning improved data quality (Diedrichs and Karplus, 1997, Weiss and Hilgenfeld, 1997). Therefore, $R_{p.i.m.}$ (precision indicating merging R factor) could be a better quality indicator, especially for the higher redundancy data. Scala was used (Evans, 2006a) to merge all diffraction data and place them on a common scale before converting the newly merged intensities to structure factors.

2.6.3 Structure refinement and model building

2.6.3.1 Phase determination by molecular replacement and calculation of electron density maps

The electron density map is calculated as a Fourier transform of the total structure factors. A structure factor is a vector with an amplitude $|F_{hkl}|$ proportional to the square root of the corresponding reflection intensity (I_{hkl}) and a phase angle (α_{hkl}). In an X-ray diffraction experiment, the structure factor amplitudes can be derived directly from the measured intensities. However, their phases are lost due to the lack of a reference waive. A number of methods exist for the solution of the “phase problem”. In the case of all structures in this study, the approximate initial phases were determined using molecular replacement. This method uses a model structure with sufficient amino acid sequence identity to the molecule in the crystal. In this case, the near-identical structure of Cdc7-Dbf4 obtained by Hughes et al. (2012) was used as a model. Prior to molecular replacement, the number of molecules in the unit cell was calculated using the Matthews coefficient (V_m) (Matthews, 1968). The Matthews coefficient uses the unit cell size and crystallographic symmetry to estimate the solvent content and number of molecules in the assymetric unit. It is a simple equation that divides the total volume of the

unit cell by the product of the molecular weight of the molecule and the number of molecules in the assymetric unit. The likely value of V_m for crystals of macromolecules covers a small range as high numbers of molecules in the unit cell will be too tightly packed, and low numbers will contain too much solvent to successfully form a crystal lattice.

Molecular replacement was performed using the programme Phaser (McCoy, 2007), which uses a maximum likelihood method to test a series of hypotheses of possible orientations of the molecule in the unit cell. Each possibility is then used to calculate a solution and the programme calculates a probability that the given data could have been observed from each solution. Phaser orients each molecule in the assymetric unit individually and in sequence. It achieves this first by sampling different rotations of the model before determining solutions for a translation function. When more than one molecule is present in the asymmetric unit, the solutions for the first molecule are tested alongside subsequently added units. The likelihood scores for given solutions diverge drastically as more molecules are added, meaning the correct solution is often quite obvious. Once this solution is determined, Phaser uses the supplied model to perform a rigid body refinement of the molecules in the asymmetric unit to improve the fit to the calculated electron density maps.

All structures solved in this study contained only one molecule in the asymmetric unit. This, in combination with the near 100% sequence identity of the model, made molecular replacement very straightforward. The XL413-bound structure obtained by Hughes et al. was used as a model for the new XL413 bound structure (PDB: 4F9C). All subsequent structures used the newly solved structure as a model.

2.6.3.2 Refinement, model building and structure validation

Refinement and model building in crystallography involves an iterative process of manual model building and automated refinement. This process improves the accuracy of the electron density maps through improvement of the calculated phases. Automated refinement interprets the electron density maps based on

knowledge of protein chemistry and the sequence of the supplied model. Observed structure factors (F_{obs}) are compared with calculated structure factors (F_{calc}), and, as the correlation of these values improves, the R_{work} value decreases. R_{work} is used as an indicator of how well the model fits the obtained data and, therefore, the likelihood that the given solution is correct. The calculated R_{work} does not factor in the inherent bias introduced by the initial model and the subsequent rounds of refinement. For this reason, the R_{free} value is calculated comparing the calculated structure factors with the observed structure factors using a small subset of data (2,000 reflections or more, typically comprising 5% of the dataset), which has not been used in refinement. When the modelling is in agreement with the data, each round of refinement will decrease both the R_{work} and R_{free} parameters. Significant divergence between R_{work} and R_{free} (> 6%) may indicate over-refinement, which can often be resolved by using tighter geometric or non-crystallographic symmetry restraints during refinement. An important consideration when using crystallographic data is the data-to-parameter ratio. Refinement of a small crystal structure (50-kDa protein) involves Cartesian coordinates and B-factors of thousands of atoms. Therefore, in macromolecular structure refinement, the data-to-parameter ratio is often very low. However, the knowledge of the physical properties of the atoms making up the unit cell helps to refine a macromolecular structure. While there is infinite number of ways atoms could be positioned in a given lattice to reproduce a given diffraction pattern, there is only one that is physically possible. Atoms do not penetrate each other and form bonds with well-known parameters, which are used as restraints when refining a macromolecular structure. At higher resolutions, the data-to-parameter ratio is improved, and restraints can be relaxed or even avoided during refinement.

For all structures refined in this study, automated refinement was performed using phenix.refine (Adams et al., 2010b). Each round of refinement produced two weighted Fourier maps with the refined phase information: *2Fo-Fc* map which should cover built parts of the model and the difference *Fo-Fc* map, which is very convenient to reveal missing or incorrectly built portions of the model. Model building was done using Coot (Emsley and Cowtan, 2004b) between rounds of automated refinement. Due to the high resolution of the collected data it was very straightforward to build newly added parts of the structure into the calculated

density. The number of residues that could be fitted correctly also increased with each round of refinement as the calculated phases improved.

Validation of the structures was performed using a range of tools in Coot, which identifies large areas of positive or negative density as well as identifying outliers in torsion angles of amino acids. Molprobity reporting, incorporated into the phenix.refine programme, also provided validation information with regard to possible clashes in the structure and geometric distortions, which allowed for fine-tuning of the structure (Chen et al., 2010). Building and validation were performed side-by-side in each round of manual model building. After each round of refinement, the backbone dihedral angles of each residue were also reviewed by plotting them on a Ramachandran plot that can be viewed in Coot. The plot shows allowed regions for the Ψ and Φ angles of each residue and identifies any residues in the structure that fall outside of the favoured regions of the plot. This is an indicator they may have been incorrectly modelled. Final structures had good geometry with over 97% of residues in the most favourite regions of Ramachandran plot.

Chapter 3. Results 1 – Design, production and activity of Cdc7-Dbf4 and Cdc7-Drf1 constructs

3.1 Aims

A crystal structure of a hetero-dimeric Cdc7 (ΔN/2q/3b)-Dbf4(MC) was previously solved in the lab (Hughes et al., 2012). Due to inherent flexibility, significant portions of the kinase were removed to assist crystallization. The deletions were made at the N-terminus of Cdc7 (residues 1-36), in KI2 (residues 228-359) and in KI3 (residues 484-529). Deletions in KI2 of Cdc7 lead to significant decreases in *in vitro* kinase activity, and it proved impossible to gain a substrate-bound structure of this construct (Hughes et al., 2012).

The aim for this part of the project was to express and purify an improved Cdc7-Dbf4 construct with activity comparable to that of the full-length kinase in quantities sufficient for crystallisation trials. In this chapter, the rationale for construct design is presented along with the production of a number of novel Cdc7-Dbf4 heterodimers with different combinations of KI2 and KI3 deletions. Numerous soluble and well-expressed constructs were obtained; revealing that only a small portion of KI2 needed to be added back to the crystallised construct to restore WT levels of activity. The construct was also further optimised to allow for crystallisation of an active heterodimer through engineering of KI3 and optimisation of different ATP analogues for the use in crystallisation trials.

3.2 Rationale for design of Cdc7 constructs with improved levels of activity

In the previously crystallised Cdc7-Dbf4 construct, deletions at the N-terminus of Cdc7 and within KI3 had no effect on the *in vitro* kinase activity. In contrast, the deletion in KI2 led to a ~2-fold reduction in activity (Hughes et al., 2012). The

truncations were designed based on sequence alignments (see appendix) and secondary structure predictions, as well as results of limited proteolysis that suggested that such regions may be disordered in the context of the dimer. The construct that yielded diffracting crystals only contained a small portion of Dbf4 (residues 210-350) corresponding to the conserved motifs-M and C with the natural linker between them. This was previously shown to be the minimal module required for Cdc7 activation, while much of the rest of Dbf4 is predicted to be disordered (Ogino et al., 2001, Kitamura et al., 2011).

Therefore, in order to obtain an active Cdc7-Dbf4 complex, residues were added back into the previously crystallised construct while maintaining the N-terminal and KI3 deletions. Residues were initially added to the Arg373 at the C-terminal end of KI2, due to its close proximity to the active site in the structure (Figure 3-1). A panel of KI2 deletions were created for expression, purification and subsequent testing in *in vitro* kinase assays. After achieving WT levels of activity, further engineering of KI3 was conducted to create mutants that would be more amenable to crystallisation (Table 3-1).

The nomenclature of deletion constructs was based on the order in which specific deletions in Cdc7 were made. The first attempted KI2 deletion was termed Δ2a and each subsequent construct was assigned the next letter of the alphabet. Upon reaching z, constructs were then named Δ2aa, 2ab and so on. The same system was utilised for KI3 deletions and the Drf1 constructs while the two MC fragments were simply named MC and MC2. This was a system developed by Siobhan Hughes that allowed simple identification of the features of each given construct.

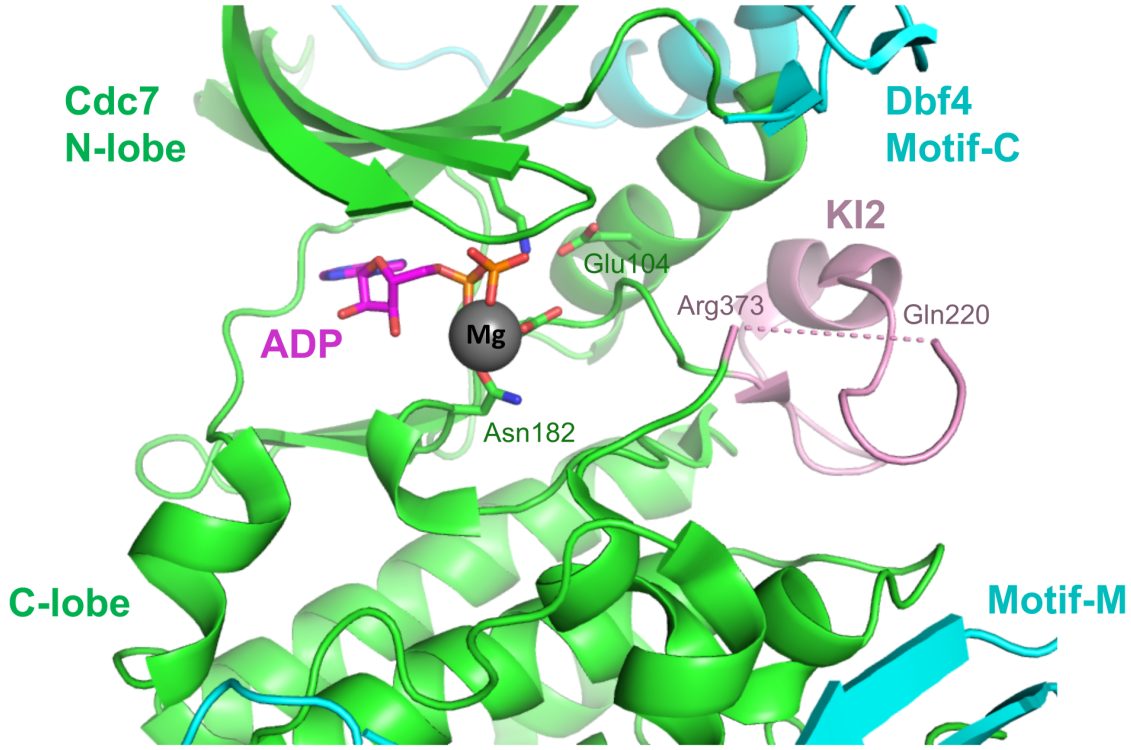


Figure 3-1 View of the Cdc7 active site in the crystal structure solved by Hughes et al. (2012)

Protein chains are shown as cartoons with conserved regions and KI2 of Cdc7 shown in green and pink respectively. Dbf4 is shown in cyan. The nucleotide (ADP, with carbon atoms in magenta) and selected Cdc7 active site residues are shown as sticks (PDB: 4F99).

Table 3-1 Cdc7, Dbf4 and Drf1 constructs produced for crystallography trials

Cdc7 Construct	ΔN	ΔKI2	ΔKI3
ΔN/2n/3b*	1-36	228-343	484-529
ΔN/2o/3b*	1-36	228-349	484-529
ΔN/2q/3b*	**1-36**	**228-359**	**484-529**
ΔN/2z/3b*	1-36	228-305	484-529
ΔN/2ad/3b*	1-36	228-282	484-529
ΔN/2ag/3b*	1-36	228-337	484-529
ΔN/2ah/3b*	1-36	228-275	484-529
ΔN/2ai/3b*	1-36	228-285	484-529
ΔN/2ak/3b	1-36	228-295	484-529
ΔN/2al/3b	1-36	231-285	484-529
ΔN/2am/3b	1-36	231-282	484-529
ΔN/2an/3b	1-36	228-288	484-529
ΔN/2ai/3c	1-36	228-285	467-AGAG-532
ΔN/2o/3c	1-36	228-349	467-AGAG-532
ΔN/2y/3c	1-36	228-265	467-AGAG-532
ΔN/2z/3c	1-36	228-305	467-AGAG-532
ΔN/2o/3d	1-36	228-349	467-AGAGA-532
ΔN/2n/3e	1-36	228-343	467-AGAG-533
ΔN/2o/3e	1-36	228-349	467-AGAG-533
ΔN/2r/3e	1-36	234-343	467-AGAG-533
ΔN/2ao/3e	1-36	234-301	467-AGAG-533
ΔN/2ap/3e	1-36	234-326	467-AGAG-533
ΔN/2aq/3e	**1-36**	**228-345**	**467-AGAG-533**
ΔN/2ar/3e	1-36	228-347	467-AGAG-533
ΔN/2as/3e	1-36	232-349	467-AGAG-533
ΔN/2at/3e	1-36	230-349	467-AGAG-533
ΔN/2o/3f	1-36	228-349	466-AGA-533

Dbf4/Drf1 Construct	Residues
Dbf4-MC*	210-350
Dbf4-MC2	214-344
Drf1-A	141-348
Drf1-B	165-348
Drf1-C	178-348
Drf1-D	211-348
Drf1-E	211-345
Drf1-F	214-345
Drf1-G	214-348

Red letters represent inserted amino acids to link ends of the truncated protein chain (*constructs made by S. Hughes). The previously crystallised Cdc7 construct and the construct crystallised in this study (ΔN/2aq/3e) are shown in bold.

3.3 Expression and purification of Cdc7-Dbf4 and Cdc7-Drf1 complexes

Recombinant heterodimers were expressed and purified as described previously (Hughes et al 2012), with some modifications. One crucial modification of the protocol was the inclusion of the phosphatase treatment, which, as explained below, resulted in considerable recovery of the activity in Cdc7-Db4 constructs harbouring complete KI2. Cdc7 constructs were expressed using the pRSFDuet-1 vector without a tag. Dbf4 and Drf1 constructs were expressed using a modified pCDFDuet-1 vector encoding an HRV14 3C protease site following the N-terminal hexahistidine tag. The vectors contain compatible origins of replications, meaning that the two plasmids are able to stably coexist in a bacterial cell. The plasmids expressing Cdc7, Dbf4 or Drf1 were introduced into *E. coli* Rosetta 2(DE3) cells and expression was induced at 16°C overnight. The proteins initially purified by batch absorption to immobilized metal affinity resin (NiNTA) were incubated with HRV14 3C protease and λ phage phosphatase to release the affinity tag and reverse autophosphorylation of the recombinant kinase. The proteins were further purified by ion exchange and size exclusion chromatography (Figure 3-2). For further details see 2.4.4.

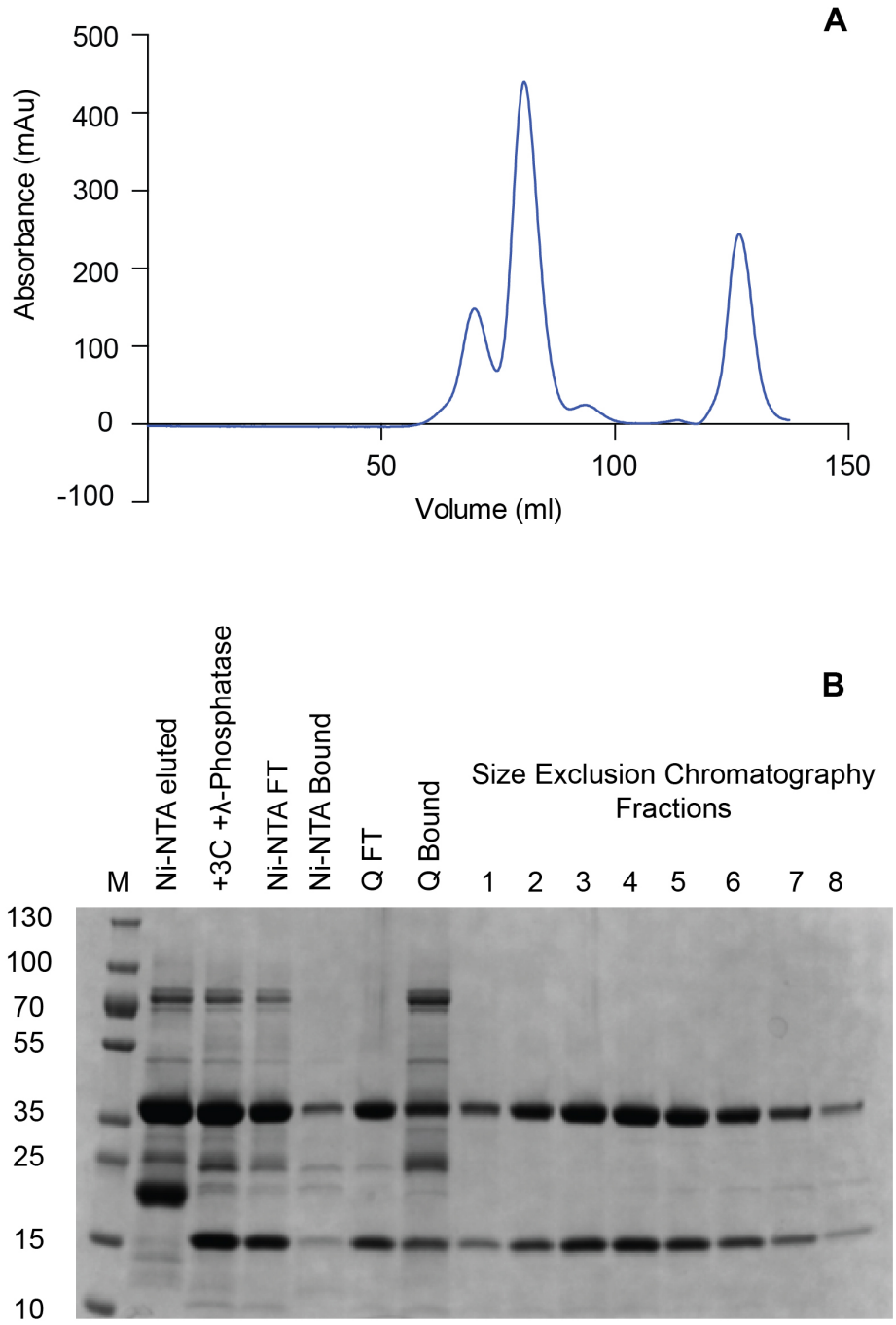


Figure 3-2 Example purification of a Cdc7-Dbf4 deletion construct

(A) Size exclusion chromatography of Cdc7(ΔN/2aq/3e)-Dbf4(MC) construct. The major peak eluting at 80 ml contained Cdc7-Dbf4. The peak was collected in 1.5-ml fractions (fractions 1-7). (B) SDS-PAGE of the various stages of the purification (Ni-NTA eluted = protein eluted from Ni-NTA beads after incubation wiith bacterial lysate, 3C+λPP=protein after cutting and dephosphorylation, NiNTA-FT and NiNTA bound are the protein fractions that passed through (low imidazole) and were then eluted (high imidazole) from the HisTrap FF column respectively, Q FT and Q bound = protein that flowed through (low salt) and was eluted (high salt) from the HiTrapQ HP column respectively. The full purification is described in section 2.4.4.

3.3.1 Dephosphorylation of recombinant Cdc7-Dbf4 leads to a substantial increase in activity

When produced in *E. coli*, active Cdc7-Dbf4 constructs undergo extensive auto-phosphorylation (Hughes et al., 2010) resulting in substantial heterogeneity of protein preparations. Furthermore, it was known that brief incubation of the kinase in the presence of ATP leads to a rapid inactivation of the enzyme (S. Hughes, unpublished observations), suggesting that autophosphorylation is likely detrimental to the activity. We reasoned that enzymatic dephosphorylation of the recombinant Cdc7-Dbf4 preparations may help improve the chances to crystallize an enzymatically active construct and allow the study of its interaction with substrates. To dephosphorylate the Cdc7-Dbf4 heterodimers, the proteins were incubated with λ-phage phosphatase in the presence of Mn^{2+}, its natural co-factor (Barik, 1993). Comparison of untreated and treated Cdc7-Dbf4 in *in vitro* kinase assays revealed a considerable difference in activities between deleted constructs and those with a full KI2. In particular, the relative activity of the previously crystallized construct Cdc7(ΔN/2q/3b)-Dbf4(MC), reduced from the ~50% reported by Hughes et al. to ~4% (Figure 3-3) upon phosphatase treatment. The kinase mutants that are inherently more active, such as those with complete KI2, are more prone to autophosphorylation during expression in bacteria, leading to gross underestimation of their activity. Therefore, the effect of the 2q deletion in the crystallized construct appears to be much greater than previously appreciated.

3.4 Optimisation of Cdc7-Dbf4 constructs for crystallography

3.4.1 Achieving wild type activity by adding residues into KI2

A number of constructs of Cdc7 were designed maintaining the deletions at the N terminus and in KI3 but increasing the number amino acids in KI2. Remarkably, addition of as few as 10 amino acid residues at the C terminus of KI2 restored or even superseded the WT levels of activity in *in vitro* kinase assays (Figure 3-3). Sequence alignments of divergent Cdc7 orthologs revealed the presence of a highly conserved Cys residue alongside four Cys residues that are invariant in metazoans (Figure 3-4).

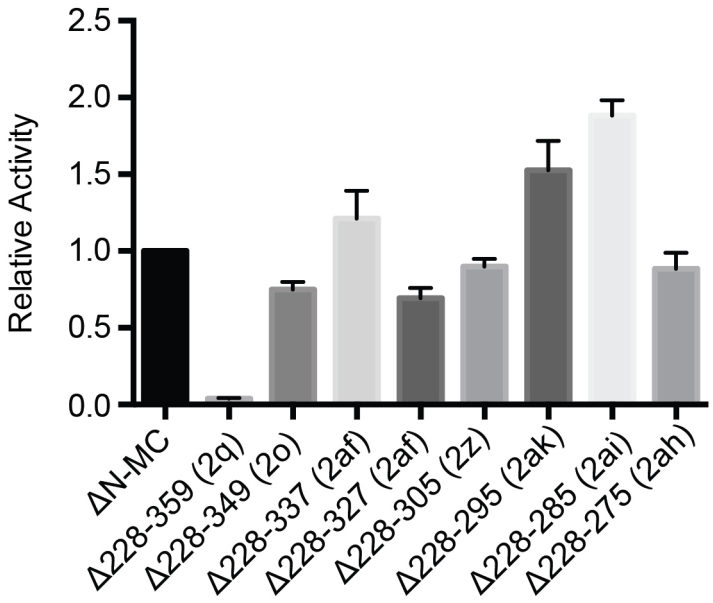


Figure 3-3 Comparison of recombinant kinase activity with different KI2 deletions

The graph shows in vitro kinase activities of Cdc7(ΔN/3b)-Dbf4(MC) constructs additionally harbouring the indicated deletions in the KI2. The activities are plotted relative to a matched construct with full KI2; error bars represent standard deviations determined from three replicates of a single experiment. ΔN/2q/3b-MC is the construct previously crystallised by Hughes et al (2012).

3.4.2 Invariant cysteine residues in KI2 are essential for kinase activity *in vitro*

In the published crystal structure of Cdc7(ΔN/2q/3b)-Dbf4(MC), two of the four invariant cysteine residues were removed as part of the 2q deletion in KI2. Results gained in this study strongly suggested that these residues could be required for the kinase activity. This was confirmed through expression and purification of recombinant Cdc7(ΔN)-Dbf4(MC) mutants. Mutations of any of the 4 invariant Cys residues lead to a drastic reduction in *in vitro* activity (Figure 3-5).

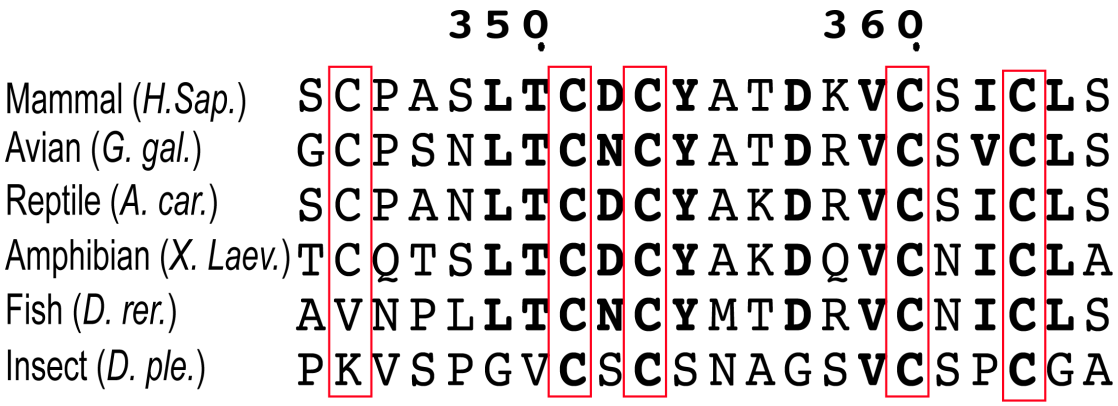


Figure 3-4 Sequence alignment of metazoan orthologues of Cdc7

Conserved and invariant residues are highlighted in red boxes.

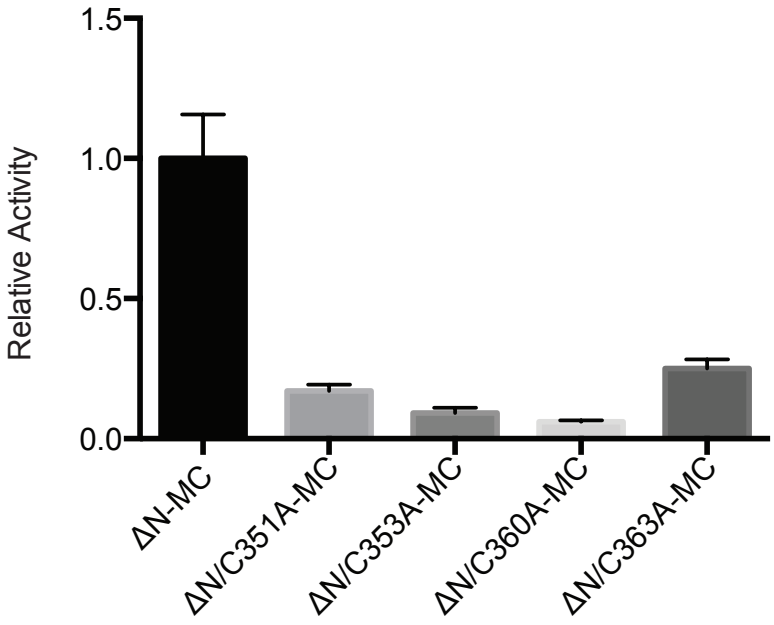


Figure 3-5 Kinase activity of mutants of invariant cysteine residues

Ala substitutions of the invariant cysteine residues Cys351, Cys353, Cys360 and Cys363 were made in the context of the Cdc7(ΔN)-Dbf4(MC) construct and tested in the *in vitro* kinase assay. The results were normalized to the activity of Cdc7(ΔN)-Dbf4(MC). Error bars represent standard deviations determined from three replicates of a single experiment.

3.4.3 Zn^{2+} increases the activity of recombinant Cdc7

The presence of clusters of invariant Cys residues is often indicative of Zn binding sites (Pace and Weerapana, 2014). Dialysis of recombinant Cdc7-Dbf4 against Kinase GF buffer (25 mM Tris pH 7.5, 150 mM NaCl) supplemented with 75 µM $ZnCl_2$ lead to a 110 ± 10% increase in activity. In contrast, the activity of a construct harbouring a disrupted tetra-cysteine motif remained almost unchanged, increasing only by ~10% **(**data not shown).

3.4.4 Optimisation of KI3

Deletions in KI3 of the previously crystallised construct had no effect on kinase activity (Hughes et al., 2012). However, in the structure of Cdc7(ΔN/2q/3b)-Dbf4(MC), 24 residues that are present in the construct were disordered and not visualised in the structure. In order to minimise this flexibility and to increase the chance of crystallization, these residues were removed from the construct and replaced with bridges of alternating Ala and Gly residues of different lengths. Optimisation of the length of the bridge allowed a complex of Cdc7-Dbf4 with a less flexible KI3 to be produced for crystallisation trials. Bridges that were too long or too short led to substantially reduced or total loss of expression of Cdc7. A number of these deletions had no detrimental effect on kinase activity *in vitro* (Figure 3-6).

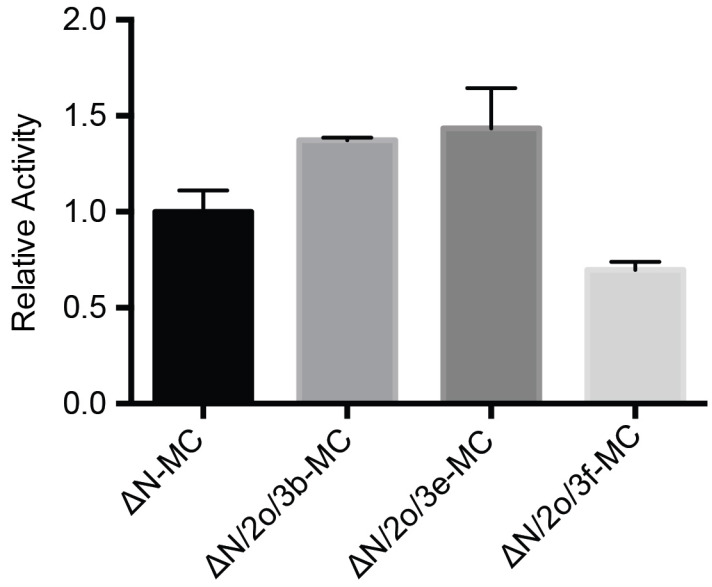


Figure 3-6 Comparison of recombinant kinase activity with different KI3 deletions

Comparison of *in vitro* kinase activities of Cdc7(ΔN/2o)-Dbf4(MC) constructs with additional deletions in KI3. Activities were normalised to ΔN-MC. Error bars represent standard deviations determined from three replicates of a single experiment.

3.4.5 Reducing the MC fragment of Dbf4 drastically reduces Dbf4 expression

As with KI3, a number of residues from the MC fragment in the structure of Cdc7(ΔN/2q/3b)-Dbf4(MC) were not visualised in the published structure, suggesting flexibility. However, reducing the size of the Dbf4 fragment by as few as 12 amino acids, as in Dbf4(MC2), lead to poor expression of the construct and a total loss of the heterodimer in purification (data not shown). For all future crystal trials, the MC fragment used by Hughes et al. was maintained.

3.4.6 Co-expression of Cdc7 with Drf1

Drf1 has also been shown to be an activating subunit when bound to Cdc7 (Morgenstern and Land, 1990, Montagnoli et al., 2002). It has been shown to play a role in early development, linking chromosome cohesion to the initiation of DNA replication in *Xenopus* egg extracts (Montagnoli et al., 2002, Takahashi and Walter,

2005a, Takahashi et al., 2008). Drf1 shares many common features with Dbf4 in that it is largely disordered with conserved motifs-N, M and C. The motifs-M and C of Drf1 share significant sequence identity with Dbf4 (45% and 53% respectively) (Figure 8.3) and therefore likely form a similar Cdc7 binding and activation module. Based on this similarity, a number of Drf1 constructs were produced guided by secondary structure predictions and sequence conservation, which covered the predicted motifs-M and C of the protein .The constructs were cloned for expression in bacteria using the modified pCDFDuet-3C vector. This allowed for successful co-expression and co-purification of the Drf1 fragments with deletion constructs of Cdc7, some of which displayed robust activity in the *in vitro* kinase assay (Figure 3-7 and 3-8).

However, despite obtaining good quantities of pure Cdc7-Drf1 heterodimers, crystallisation trials failed with all constructs using a number of different nucleotide analogues. This suggests that Drf1 may have a different effect on the orientation of the N and C lobe of Cdc7 to Dbf4, preventing the protein from forming suitable crystal contacts. Interestingly, adding residues beyond the conserved M and C motifs of Drf1 lead to significant decreases in activity. It is not possible to say whether this is due to the stability of the purified complex or whether there is an inhibitory role for the regions of Drf1 outside of the conserved activation motifs.

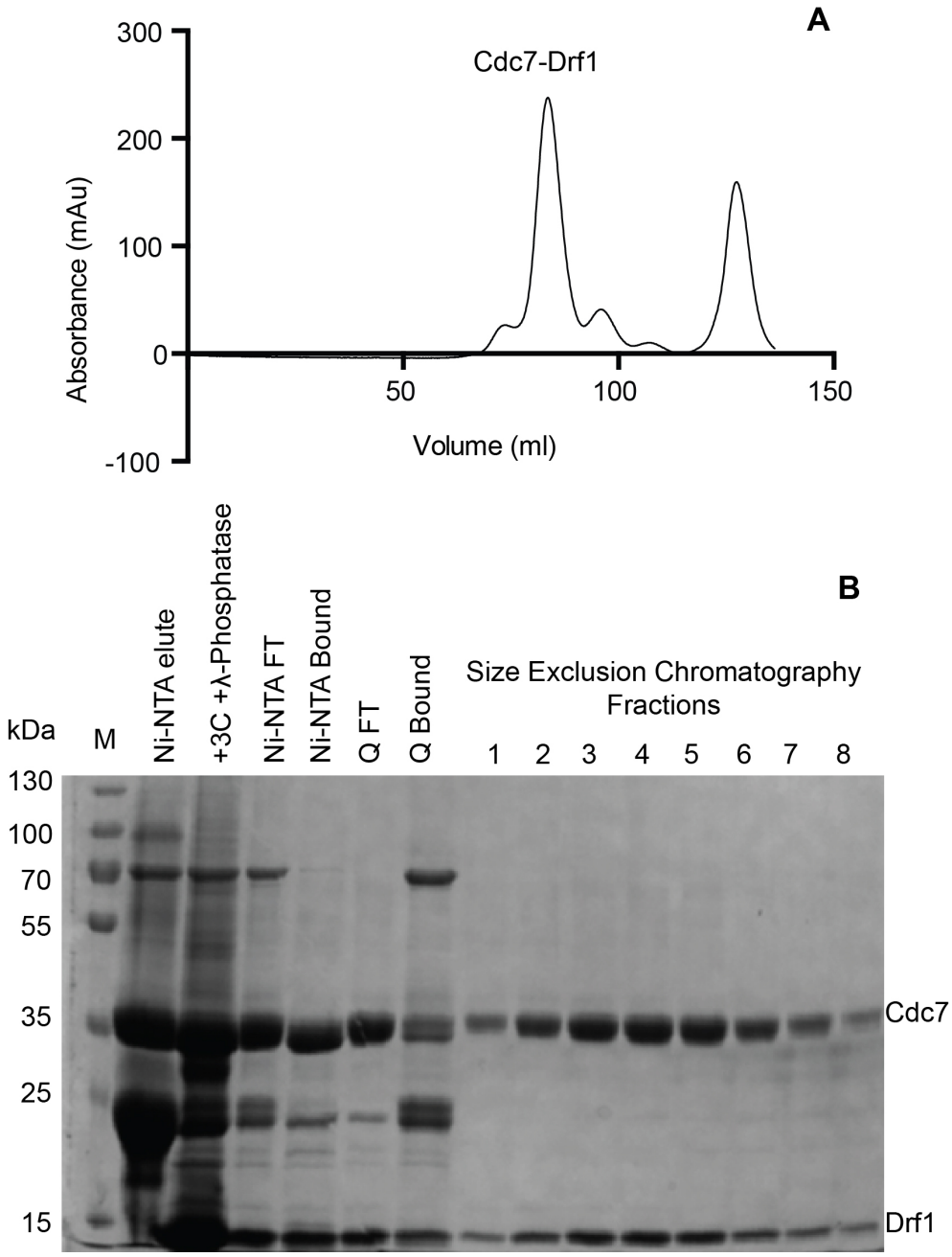


Figure 3-7 Example purification of the Cdc7(ΔN/2aq/3e)-Drf1 (211-345) construct

(A) Size exclusion chromatography; the major peak eluting at ~85 ml contained the recombinant heterodimer. The peak was collected in 1.5-ml fractions (fractions 1-7). (B) SDS-PAGE of the various stages of the purification. Ni-NTA eluted = protein eluted from Ni-NTA beads after incubation with bacterial lysate, 3C+λPP=protein after cleavage of the His_6 tag and dephosphorylation, NiNTA-FT and NiNTA bound are the protein fractions that passed through (low imidazole) and were then eluted (high imidazole) from the HisTrap FF column respectively, Q FT and Q bound = protein that flowed through (low salt) and was eluted (high salt) from the HiTrapQ HP column respectively. The full purification is described in section 2.4.4.

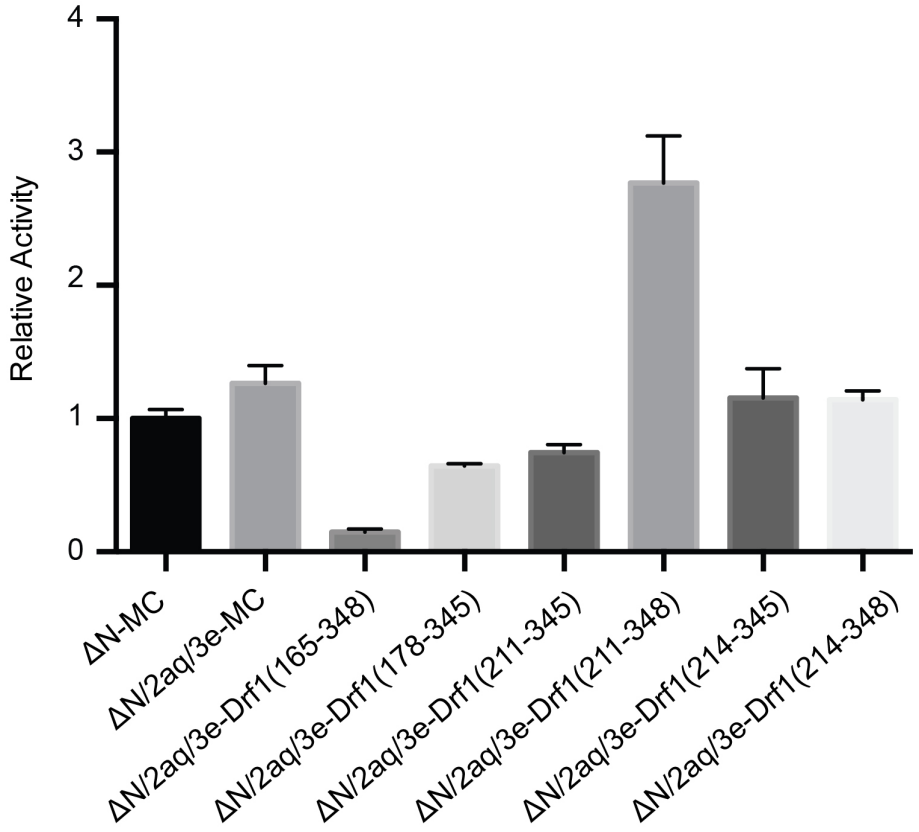


Figure 3-8 Activities of Cdc7(ΔN/2aq/3e) bound fragments of Drf1 compared to Cdc7(ΔN)-Dbf4(MC) and Cdc7(ΔN/2aq/3e)-Dbf4(MC)

Activity normalised to Cdc7(ΔN)-Dbf4(MC). Error bars represent standard deviations determined from three replicates of a single experiment.

3.4.7 Optimisation of bound nucleotide

Binding of ligands to the active site is expected to induce subtle changes to the relative orientations between the N- and C-lobes of the kinase, which may help in crystallization. For this reason, a number of different nucleotide analogues and Cdc7 inhibitors were tested in crystallisation trials with a number of promising constructs (Table 3-2). Nucleotide analogues were used in the range of 1-5 mM either alone or in the presence of $MgCl_2$ (2.5-10 mM).

Table 3-2 Table of nucleotide analogues and Cdc7 inhibitors used in crystallography trials

Ligand	Type	Supplier
AMP-PNP	Nucleotide Analogue	Jena Bioscience
ATPγS	Nucleotide Analogue	Sigma-Aldrich
AMP-PCP	Nucleotide Analogue	Jena Bioscience
ADP-AlF_4^-	Nucleotide Analogue	Sigma-Aldrich
ADP-BeF_3^-	Nucleotide Analogue	Sigma-Aldrich
XL413	Cdc7 Inhibitor	Sigma-AldrichXlysis
PHA767491	Cdc7/Cdk9 Inhibitor	Sigma-Aldrich
CRT	Cdc7 Inhibitor	Cancer Research Technologies

3.5 Conclusions

A total of 26 deletion constructs of Cdc7 were expressed and purified in various combinations with two Dbf4 and seven Drf1 fragments. A number of constructs were produced in sufficient quantities for crystallography with activity equivalent to that of a WT kinase in *in vitro* kinase assays. Addition of as few as 10 amino acids restored full kinase activity and examination of this region of the protein revealed a potential metal binding site composed of 4 invariant Cys residues, which was disrupted in the previously crystallised construct. A significant increase in *in vitro* activity in the presence of Zn salts further supported the idea that metal binding may be important in correct folding of the more complete construct. Despite achieving full activity with small changes to KI2, significant engineering of KI3 was required to reduce the flexibility of the kinase for crystallisation. Further deletions in this insert were not detrimental to kinase activity and the heterodimeric complexes were sufficiently stable to allow co-purification from bacteria. Attempts to reduce the flexibility of the complex further by deletion of disordered residues in the MC fragment of Dbf4 were unsuccessful leading to a significant reduction in kinase yields during purification.

Increasing the size of the Drf1 fragment beyond the conserved motifs-M and C also lead to a significant decrease in activity (Figure 3-8). It is impossible to say without structural data whether this is due to the complex being unstable or whether the less ordered regions of Drf1 have auto inhibitory effects. A lack of progress in attempting to crystallise Cdc7-Drf1 suggests that the binding mode of Drf1 is sufficiently different to prevent crystallisation. This may be due to properties of Drf1 itself or the orientation of the N and C lobes of Cdc7 when Drf1 is bound. Further structural work on this complex may reveal important features for Cdc7 activation and could explain the different roles of Cdc7 bound to each activating partner.

It is interesting that high levels of activity can be achieved *in vitro* with deletions of residues 228-345 and 467-533 in KI2 and KI3, respectively. Therefore, these portions of the inserts play no role in the minimal peptide substrate binding and catalysis. Conceivably, they could be involved in protein-protein interactions required for Cdc7 to fulfil its many roles in the cell cycle. Further functional work *in cellula* will be carried out to ascertain the importance of these inserts in a more physiological setting (Chapter 6). Further to this, the clear role that phosphorylation plays in the activity of Cdc7 *in vitro* is indicative of another layer of Cdc7 regulation beyond Dbf4 and Drf1 binding.

The following chapters build on the production and characterisation of the various Cdc7-Dbf4 complexes presented here, providing structural data on an active Cdc7-Dbf4 construct bound to an ATP analogue and a potent Cdc7 Inhibitor in chapter 4. In chapter 5 the requirements of substrate binding are revealed by the structure of a Cdc7-Dbf4 construct bound to an Mcm2-derived substrate peptide.

Chapter 4. Results 2 – Crystallisation of a minimal Cdc7-Dbf4 construct bound to XL413 and ADP-BeF_3^-

4.1 Aims

The main aim of this project was to determine a crystal structure of a deletion construct of Cdc7-Dbf4 with improved activity compared to the strongly attenuated previously crystallised construct. The expectation was that the structure would reveal novel features that are critical for the kinase activity. As described in the previous chapter, a number of different Cdc7 constructs were expressed and purified with different mutants of the activating subunits Dbf4 and Drf1. This construct screening approach resulted in successful crystallization and structure determination of one highly active heterodimeric construct.

In this chapter, the structures of Cdc7(ΔN/2aq/3e)-Dbf4(MC) bound to a potent Cdc7 inhibitor XL413 and a non-hydrolysable ATP mimic ADP-BeF_3^- are presented. The inhibitor bound structure was refined to a resolution of 1.44 Å. As previously observed by Hughes et al., the small molecule is bound in the active site of the kinase. The structure reveals a previously unseen Zn binding domain in KI2 of Cdc7. This domain pins back the activation loop of Cdc7 against motif-M of Dbf4, sandwiching it against the C-lobe of the kinase. This appears to be vital for proper engagement of the kinase with Dbf4, resulting in an open platform for substrate binding. By superimposing previously obtained kinase structures with the new structure of Cdc7 it is possible to hypothesise residues that may be involved in substrate binding, explaining the minimal consensus sequence of Cdc7.

The structure of Cdc7(ΔN/2aq/3e)-Dbf4(MC) bound to ADP-BeF_3^- was obtained and has the same overall architecture as the XL413 bound kinase. However, the structure also reveals the position of ATP γ-phosphate, mimicked by the BeF_3 moiety, as well as the pair of Mg^{2+} cation cofactors, which are essential for the catalysis of the phosphoryl transfer.

4.2 Crystal screens, optimisation and freezing of Cdc7(ΔN/2aq/3e)-Dbf4(MC)^{XL413}

4.2.1 Initial screening of kinase complexes

Crystal screens were performed with each expressed and purified kinase complex in the presence of various ATP analogues or small molecule inhibitors (Table 3-2), in the presence or absence of $MgCl_2$ chloride (2.5-20 mM). Kinase complexes, concentrated to approximately 10 mg/ml were incubated with the ligands for 10 min on ice prior to crystallization screening. High-throughput screens were setup with each purified protein at 18°C in the order presented (Table 2.1) using 200-nl drops aliquoted by a Mosquito robot (TTP LabTech). For each construct, as many sparse matrix conditions were trialled as possible with the amount of protein obtained. A typical experiment used ten to twelve 96-well crystallization trays (960 to 1152 sparse matrix conditions). Initial crystals were obtained with Cdc7(ΔN/2o/3e)-Dbf4(MC)$^{AMP\text{-}PNP}$. However, these diffracted poorly even after painstaking optimisation. Eventually diffracting crystals were obtained with the development of the Cdc7(ΔN/2aq/3e)-Dbf4(MC) construct.

4.2.2 Crystallisation of Cdc7(ΔN/2o/3e)-Dbf4(MC)$^{AMP\text{-}PNP}$

A minimal construct Cdc7(ΔN/2o/3e)-Dbf4 (MC), was successfully crystallised in the presence of AMP-PNP in a number of conditions containing PEG 8000. These conditions produced rod-like crystals (Figure 4.1). Despite extensive optimisation of the precipitant concentration, buffer pH, $MgCl_2$ and NaCl concentration, these crystals only diffracted to ~13 Å on a synchrotron X-ray source, and could not be optimised further.

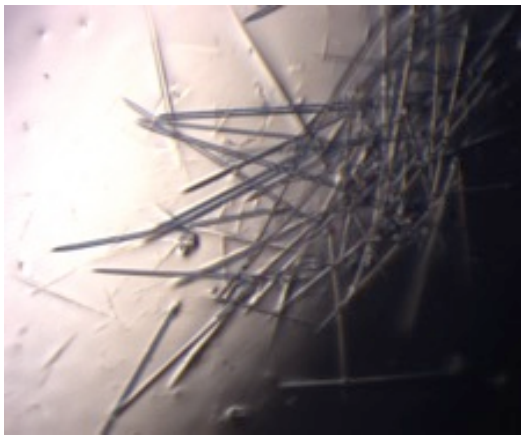


Figure 4-1 Crystals of Cdc7(ΔN/2o/3e)-Dbf4 (MC) grown in the presence of AMP-PNP

Crystals were grown at 18°C by vapour diffusion in hanging drops.

4.2.3 Optimisation of Cdc7(ΔN/2aq/3e)-Dbf4(MC)^{XL413} crystallisation

Addition of four more amino acids back into KI2 gave the Cdc7(ΔN/2aq/3e)-Dbf4 (MC) construct, which in the presence of XL413 readily yielded polycrystalline precipitates in a number of conditions with acidic pH. Numerous batches of the heterodimeric construct were used to optimise buffer pH, precipitant, NaCl and $ZnCl_2$ concentration at 18°C and 4°C. Further to this, additive and ionic liquid screens were utilised to gain crystals of sufficient size and quality for X-ray diffraction (Figure 4.2). The best crystals grew in 0.1 M MIB (sodium malonate, imidazole, boric acid) buffer pH 5.5, 21% poly ethylene glycol (PEG) 1500 and

5mM $MgCl_2$. Crystals were grown in manually created 2 µl hanging drops at 18°C and generally grew to full size within 16 h.

Figure 4-2 Crystals of Cdc7(ΔN/2aq/3e)-Dbf4 (MC) grown in the presence of XL413

Crystals were grown by vapour diffusion in hanging drops in 0.1 M MIB buffer pH 5.5, 21% PEG 1500 and 5 mM $MgCl_2$. Crystals were grown at 18°C and harvested after 48 hours.

4.2.4 Optimisation of cryo conditions

Crystals grown in optimised conditions diffracted to approximately 1.8 Å resolution on a Rigaku MicroMax-007 generator, equipped with an image plate detector. Prior to data collection at a synchrotron, snap freezing conditions were optimised to avoid degradation of X-ray diffraction due to crystal cracking or shrinkage caused by cryo-protectants. Crystals, harvested from mother liquor were incubated in 5-ml drops containing 2.5 mM XL413, 100 mM NaCl, 2.5 mM $MgCl_2$, 10 mM Tris-HCl and 80 mM MIB, pH5.5, with stepwise increases in PEG 1500 (21-25% in 1% steps) and the cryo-protectant PEG 400 (5-25% in 5% steps). An optimal freezing condition was achieved at 24% PEG1500 and 20% PEG 400.

4.3 Data collection and structure refinement of Cdc7(ΔN/2aq/3e)-Dbf4(MC)^{XL413}

4.3.1 Data collection

Crystals of Cdc7(ΔN/2aq/3e)-Dbf4(MC)^{XL413} complex grown and cryopreserved under optimized conditions were used for X-ray data collection on the BM14 beamline of the European Synchrotron Radiation Facility (ESRF, Grenoble, France). BM14 is a macrocmolecular crystallography beamline with a tuneable energy range of 7-17 keV. The beamline is also equipped with a MarCCD detector with an active area of 225x225 mm. The majority of the crystals diffracted to 1.3-2.5 Å resolution, and the best dataset with measurable signal extending to 1.44 Å was used to determine and refine the structure (Figure 4-3, Table 4-2).

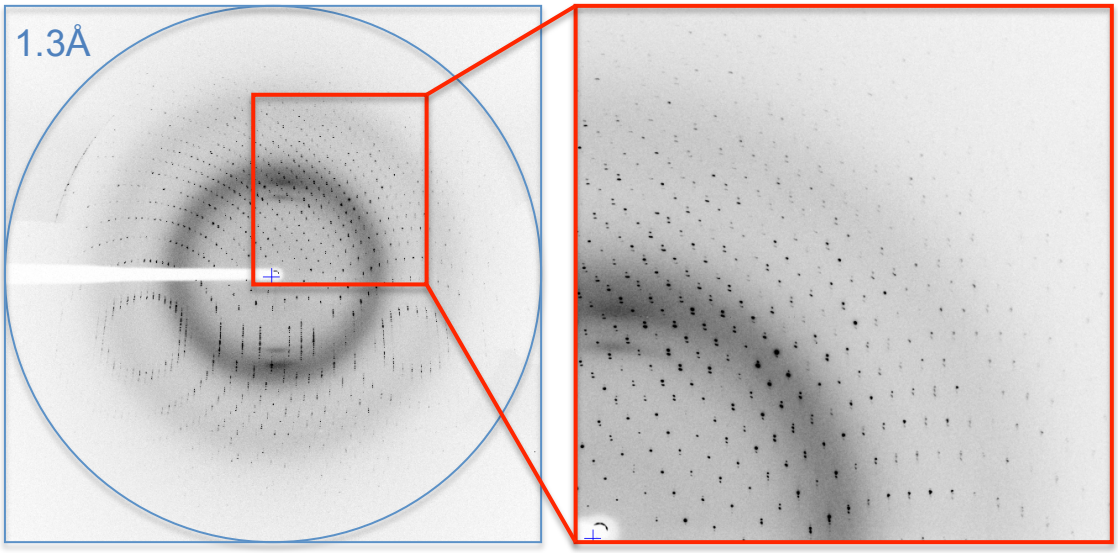


Figure 4-3 A sample X-ray diffraction pattern from a crystal of Cdc7(ΔN/2aq/3e)-Dbf4 (MC)^{XL413}

4.3.2 Data processing and structure refinement

The collected X-ray diffraction data were processed using XDS (Kabsch, 2010b) and Scala (Evans, 2006b) via the Xia2 automatic data reduction pipeline (Winter, 2010b). The best dataset contained measurable signal (<*I*/sig*I*> over 1.5) to the highest resolution shell of 1.491-1.440 Å. The structure was determined by molecular replacement using Phaser (McCoy et al., 2007), and built using Phenix Autobuild (Adams et al., 2010a). The previous crystal structure of Cdc7-Dbf4 was used as a search model (PDB: 4F9C). The structure was refined using phenix.refine (Adams et al., 2010a) and Refmac (Murshudov et al., 1997) with manual model building in Coot (Emsley and Cowtan, 2004a). X-ray data collection and refinement statistics are given in Table 4-1.

Table 4-1 Data collection and refinement statistics for Cdc7(ΔN/2aq/3e)-Dbf4(MC) bound to XL413 and ADP-BeF_3

	XL-413	**ADP-BeF_3^-**
Data collection *		
Beamline	ESRF BM14	DLS I03
Space group	P 41 21 2	P 41 21 2
Unit cell parameters		
a, b, c (Å)	61.64, 61.64, 233.91	61.15, 61.15, 235.44
α, β, γ (^{o})	90, 90, 90	90 90 90
Wavelength (Å)	0.9796	0.9796
Resolution range (Å)	42.85 -1.44 (1.491 -1.44)	48.24 - 1.79 (1.854 - 1.79)
Average multiplicity	9.1 (6.7)	11.5 (7.4)
Completeness (%)	99.52 (95.76)	98.98 (91.15)
CC1/2	1 (0.789)	0.999 (0.827)
CC*	1 (0.939)	1 (0.951)
R_{merge}	0.05269 (0.6114)	0.0698 (0.6182)
R_{meas}	0.05585 (0.6621)	0.07311 (0,6639
</>/<sig(*I*)>	26.48 (2.96)	23.66 (3.02)
Wilson <B> (Å^2)	13.55	21.62
Solvent content (%)	47	47
Refinement		
Resolution	**42.85 -1.44**	**48.24 - 1.79**
Reflections (total/free)	750697/52186	492798/42930
R_{work}/R_{free}	0.1649/0.1969	0.1877/0.2192
CC (work)	0.969 (0.852)	0.966 (0.880)
CC (free)	0.951 (0.840)	0.956 (0.881)
No. Atoms	4022	3718
Protein	3534	3535
Ligand/ion	38	35
Water	450	148
Average B-factors (Å^2)	20.70	24.50
Protein	19.30	24.50
Ligand/ion	28-20	17.60
Water	31.00	26.00
R.m.s. deviations from ideal		
Bond lengths (Å)	0.013	0.007
Bond angles (°)	1.35	1.15
Ramachandran (%)		
Favoured	97	97
Outliers	0.27	0.27
Number of TLS groups	6	6

*Data for the highest resolution shell are shone in parentheses.

4.4 The structure of Cdc7(ΔN/2aq/3e)-Dbf4(MC)^{XL413}

4.4.1 A novel zinc binding domain is present in KI2

Cdc7(ΔN/2aq/3e)-Dbf4(MC)^{XL413} crystallised in the space group $P4_12_12$ with one molecule of the kinase present in the asymmetric unit. The newly added residues of Cdc7, shown in orange in Figure 4-4, comprise a previously unseen portion of the activation loop in KI2. The crystal structure reveals the presence of a novel Zn binding domain within KI2 (Figure 4-4). Within this domain, Cys351, Cys353, Cys360 and Cys363 coordinate a single metal atom, in excellent agreement with the results of the mutagenesis analyses (Figure 3-5). The presence of the Zinc atom in the structure was confirmed by anomalous X-ray scattering using X-ray diffraction data collected at a wavelength of 1.2837 Å, corresponding to the absorption K-edge of Zn (Figure 4-5).

Zinc binding domains are found across a range of proteins and are particularly common in proteins involved in DNA binding proteins such as those required for transcription. Zinc binding folds can be classified into a number of groups based on the sequence homology and structural features involved in the coordination of the Zinc ion (Krishna et al., 2003). Zinc binding domains have been identified in a number of kinases. Zinc ribbon folds, one group of zinc binding domains have been identified in kinases, including adenylate kinase from *Bacillus stearothermophilus* (PDB:1ZIN) in which the zinc binding domain is found in the lid region of the active site and plays an integral role in stabilizing the structure of the kinase (Berry and Phillips, 1998). The zinc ribbon fold is also seen in the Tec family of Tyrsoine kinases in which the zinc is found in a Btk motif which is present downstream of the pleckstrin homology domain of the protein (PDB :1B55) (Baraldi et al., 1999). Zinc binding has also been shown to play a regulatory function in kinases, Protein kinase C (PKC) contains zinc a zinc binding domain which is able to bind to lipid second messengers, binding of which leads to activation of the kinase. PKC can also be activated by changes in redox state and work by Zhao et al. showed that this activation is due to a common mechanism in which lipid messenger binding leads to Zinc release and significant structural changes. This change in structure is

integral to kinase activation under such conditions (Hubbard et al., 1991, Zhao et al., 2011). The Zinc binding domain identified in Cdc7 appears to most closely resemble the short zinc binding loop class of folds. Such folds are generally small but separate domains which are stabilized by the binding of Zinc. They are not typically part of larger secondary structural elements, though in some cases one coordinating residue may be part of a secondary structural feature. Such a fold is present in the atypical kinase structure of a transient receptor potential (TRP) channel which has retained its phosphotransferase activity (Yamaguchi et al., 2001). Another interesting example of a Zinc binding domain in a kinase structure is that of casein kinase 2 (CK2), the most closely related kinase to Cdc7. The domain present in CK2 is similar to that of Cdc7 in that it involves four invariant cysteine residues, which coordinate the zinc in a tetrahedral conformation. However, in CK2 the Zinc binding domains of two individual monomers form extensive interactions with each other, which facilitates dimerization (Chantalat et al., 1999). This represents an intriguing possibility that the Zinc binding domain of Cdc7 may be involved in such a protein-protein interaction that is important for its function in the cell. However it is not possible to ascertain this from the obtained structure due to the significant portion of KI2 that has been removed which may occlude such an interaction in the context of the full protein. Despite the presence of Zinc binding domains in kinase structures, the domain observed in Cdc7 appears to be the first example, which is present in the activation loop of the protein. This may be due to the significant length of the activation loop of Cdc7 relative to other kinases, requiring structural features such as a Zinc binding domain to maintain a structure suitable for activity.

4.4.2 The Cdc7 KI2 Zn binding domain interacts with motif-M of Dbf4 resulting in full ordering of the activation segment

In the published structure the activation segment was not fully ordered (Figure 3-1), and the active site seemed to be partially occluded by a slight movement of Cdc7 Arg373. In the new structure, the KI2 Zn binding domain tightly interacts with Dbf4 motif-M, thereby pinning the activation segment of Cdc7 to the C-lobe. Unexpectedly, in the new structure Dbf4 motif-M is sandwiched between the Zn-binding domain of KI2 and the C lobe of the kinase. This arrangement may help to

ensure that full activation of the kinase takes place only upon proper engagement of Dbf4 motifs-M and -C.

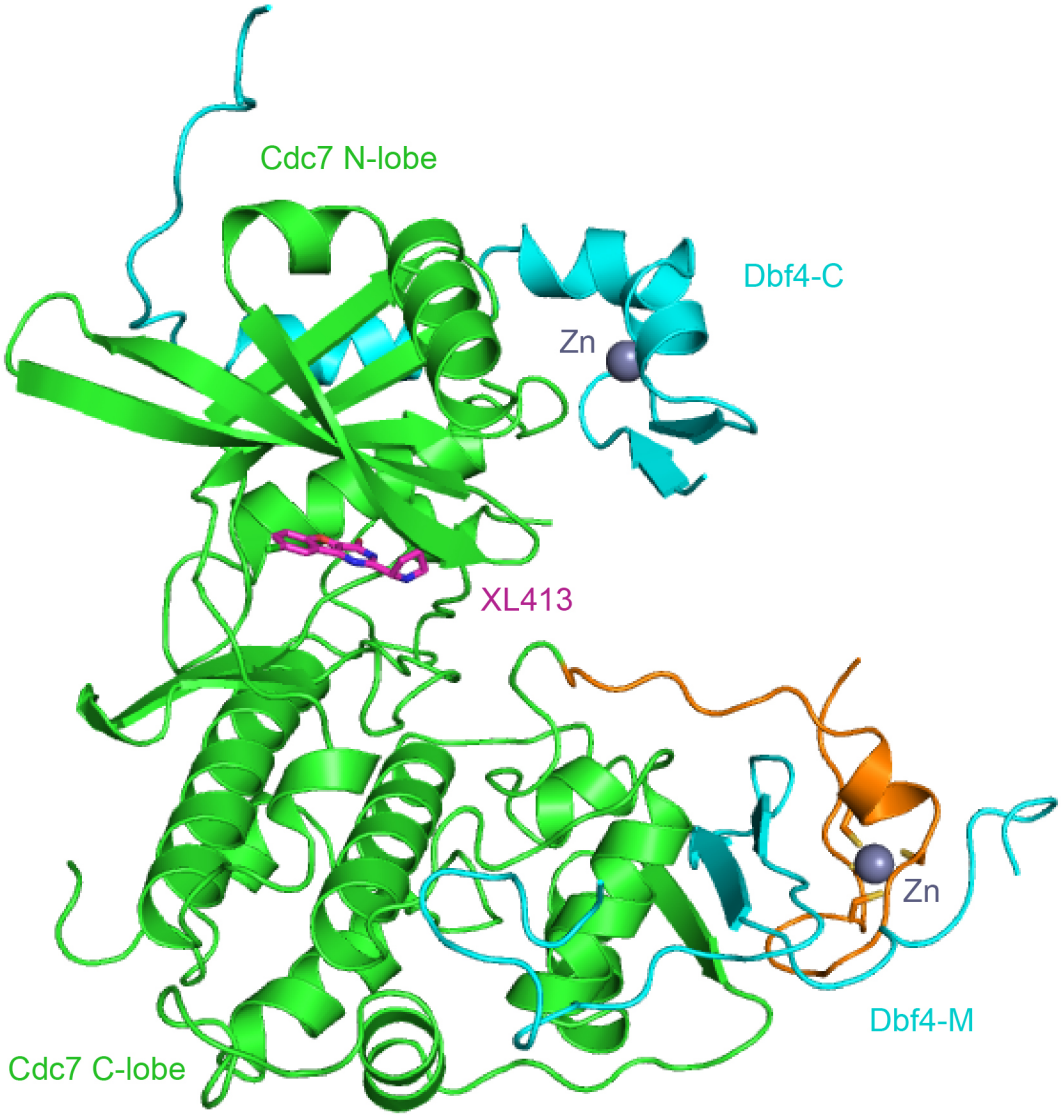


Figure 4-4 Overall structure of the ΔN/2aq/3e-MC construct of Cdc7-Dbf4 bound to XL413

Protein chains are shown as cartoons with Dbf4 shown in cyan. Cdc7 residues present in the structure by *Hughes et al* are shown in green and the newly added residues are shown in orange. XL413 is shown in pink as sticks with non-carbon atoms following standard colouration. Zinc atoms are shown as grey spheres.

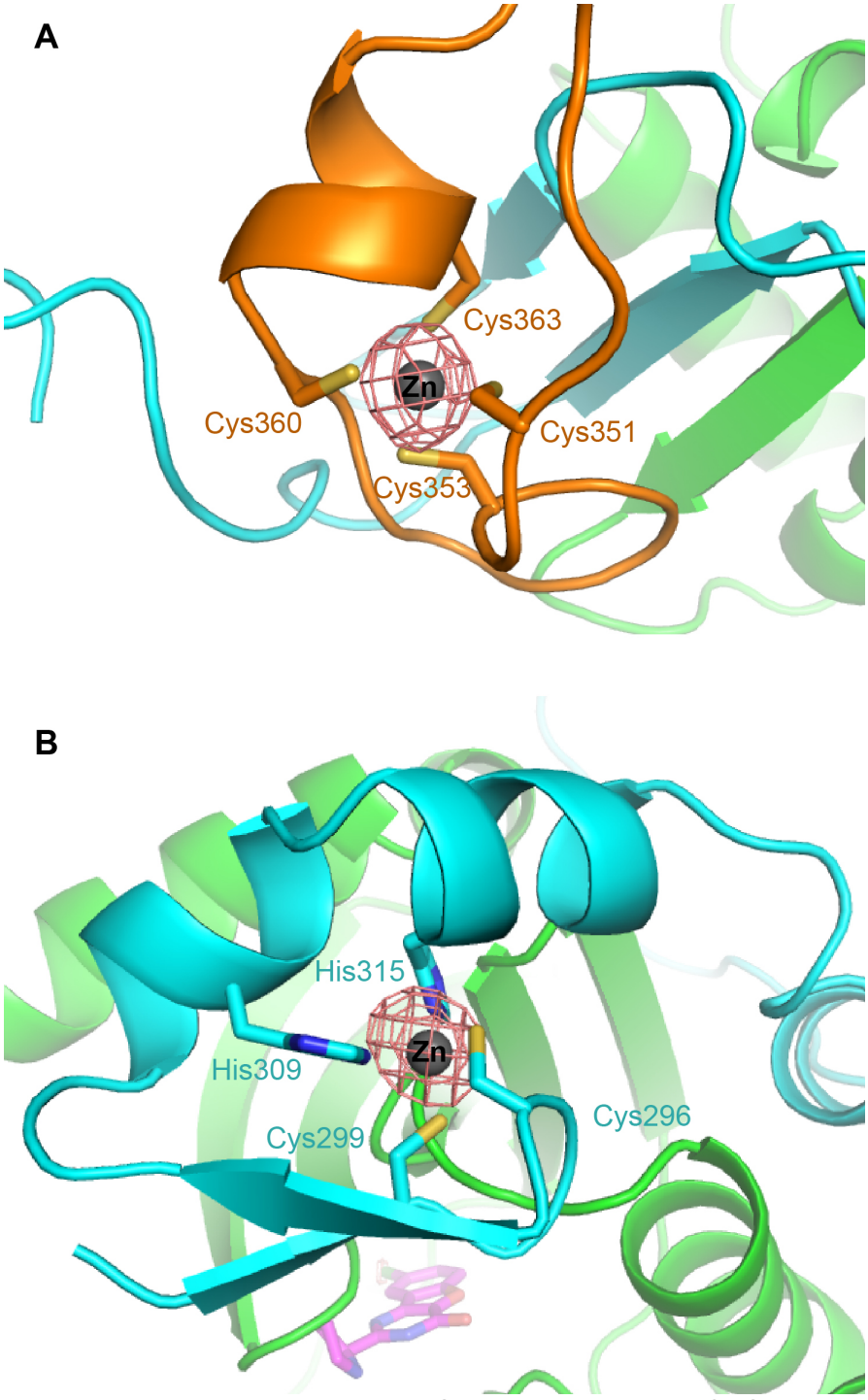


Figure 4-5 Zinc binding domains in Cdc7 Kl2 (A) and Dbf4-C (B)

Metal binding His and Cys residues are shown as sticks. Phased anomalous electron density maps are shown as pink chicken wire, contoured at 5.0σ.

4.4.3 Comparison of ΔN/2aq/3e-MC to the previous structure

Superposing the bound crystal structure of Hughes et al (4F9C, bound to XL413) with the newly obtained structure shows results in r.m.s. displacements for the common Cα atoms of only 0.615 Å (Figure 4-6). The largest change is brought about by the presence of the KI2 Zn-binding domain. Close inspection of the structures confirms the suspected reason for the reduced activity of the previously crystallised construct. The partially disordered activation loop projects into the cleft of the kinase, likely partially obstructing access to the active site (Figure 4-7). The Zn-binding domain pins the activation loop, removing this portion of the activation loop from the active site allowing for substrates to enter the active site.

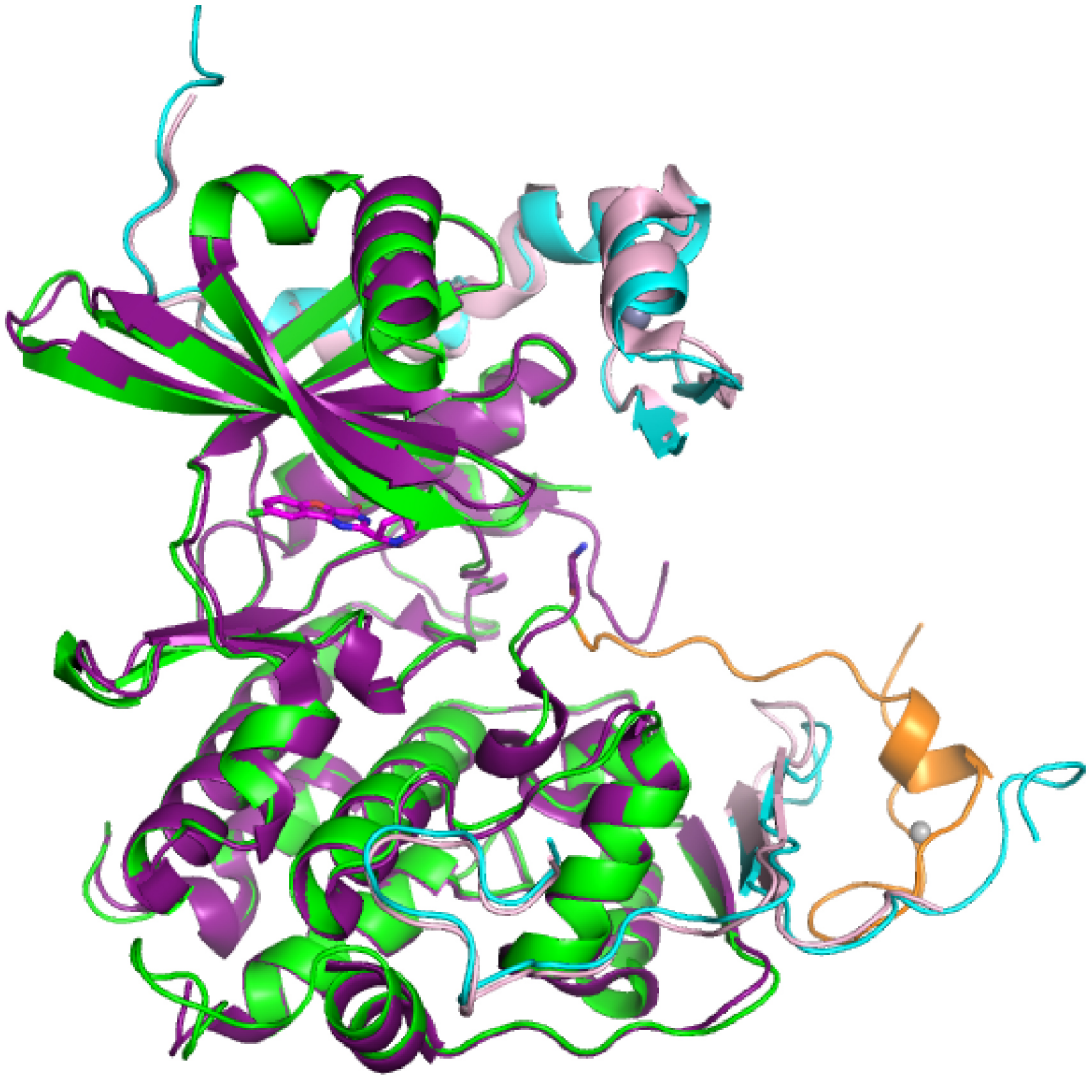


Figure 4-6 Overall structure of the ΔN/2aq/3e-MC construct of Cdc7-Dbf4 superimposed on the previously crystallised structure

Cdc7 and Dbf4 of the structure solved by Hughes et al. (2012) (PDB 4F9C) is shown in purple and pink for Cdc7 and Dbf4 respectively.

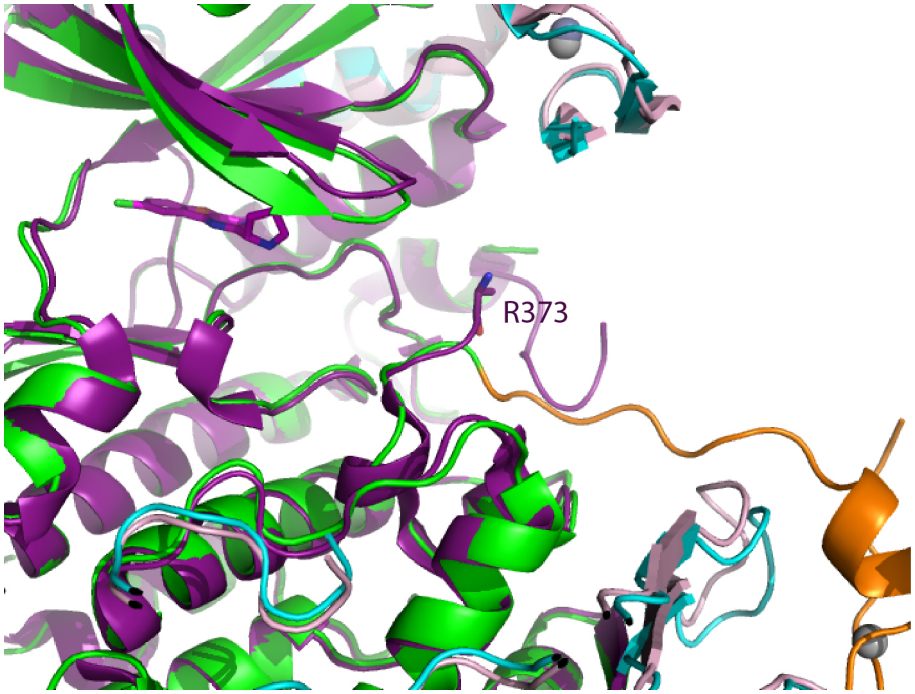


Figure 4-7 Close up of the active site of Cdc7(ΔN/2aq/3e)-Dbf4(MC) superposed on the published structure.

Residues discussed in the text are shown as sticks with non-carbon atoms following standard colouration.

4.4.4 Comparison with a peptide bound PKA structure highlights Cdc7 residues potentially involved in substrate binding

Superposition of a substrate peptide-loaded structure of PKA (PDB: 3X2W) with Cdc7(ΔN/2aq/3e)-Dbf4(MC) suggested that Arg373 and Arg380 may be poised to make interactions with the P+1 position of the Cdc7 substrate peptide, normally occupied by an acidic residue of a pre-phosphorylated Ser (Cho et al., 2006) (Figure 4-8). Intriguingly, both Arg373 and Arg380 are invariant among Cdc7 orthologues. To confirm their importance for Cdc7 function, the recombinant Cdc7-Dbf4 heterodimers harbouring substitutions at these positions were produced. *In vitro* kinase assays revealed a significant decrease of activity associated with the mutations (Figure 4-9). This decrease in activity was observed with peptides containing sites containing a pre-phosphorylated serine at the N+1 position and a sit containing an SE site. However, no further requirements for substrate binding

can be inferred from the current structure, highlighting the need for crystallization of a substrate-bound form of the kinase.

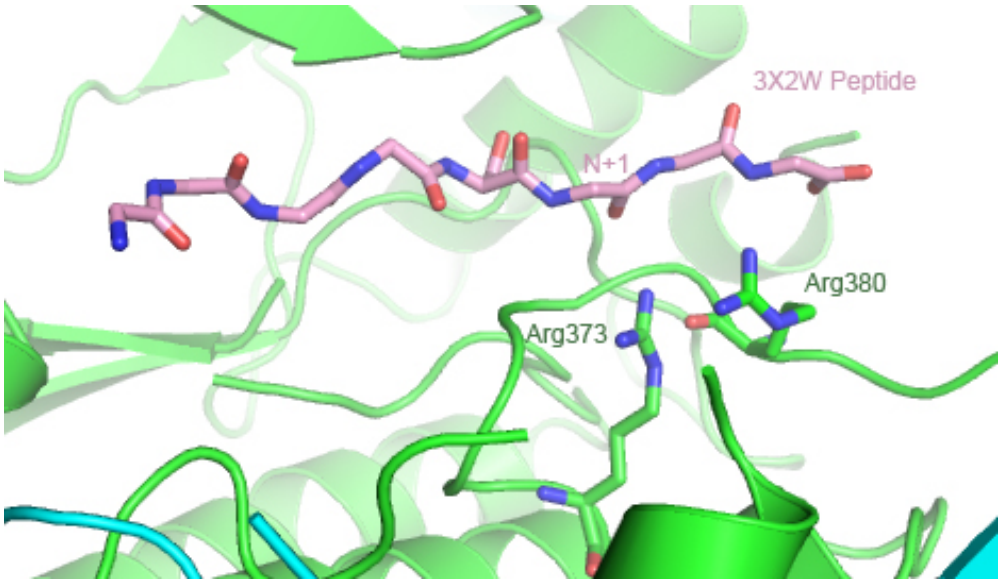


Figure 4-8 A model for Cdc7-substrate interaction

The model was constructed by superposition of a substrate peptide-bound PKA structure (PDB: 3X2W) with the (ΔN/2aq/3e)-Dbf4 (MC) structure. Cdc7-Dbf4 is shown as cartoons, and the backbone of the substrate peptide from 3X2W is represented as sticks. For clarity, the chain of PKA is hidden.

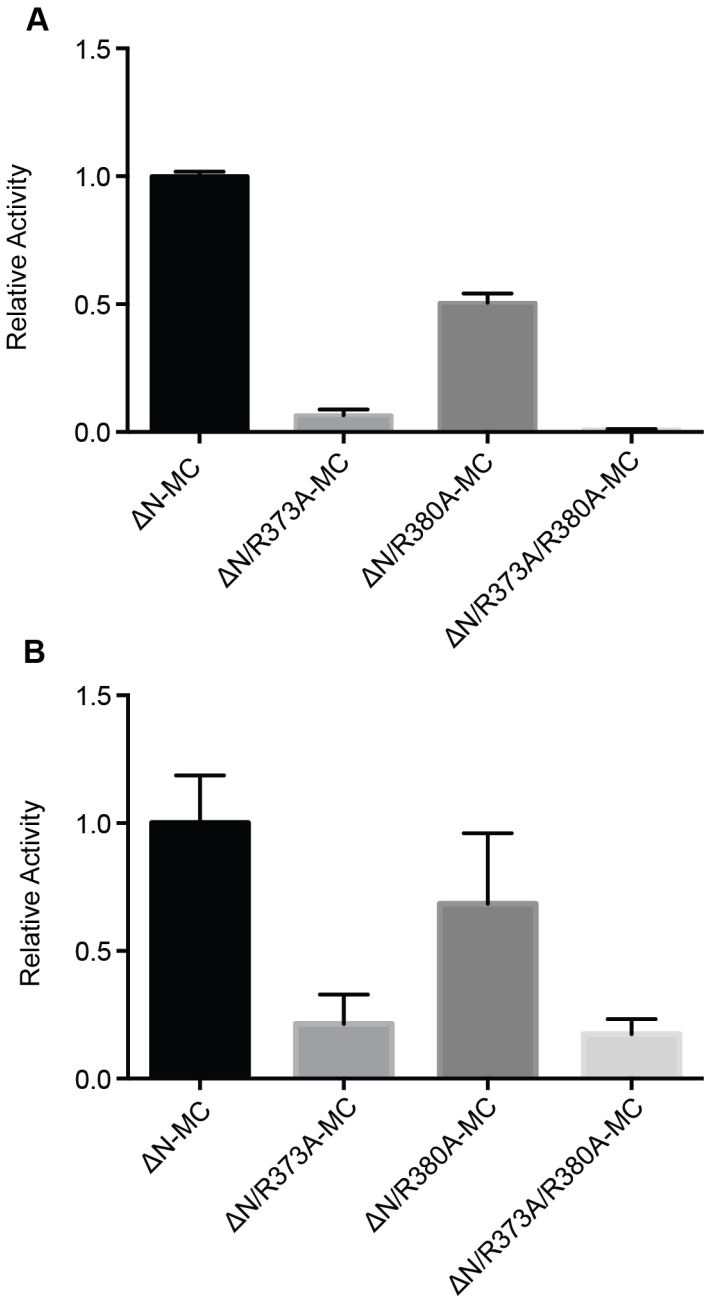


Figure 4-9 Arg373 and Arg380 are essential for Cdc7 activity *in vitro*

Alanine substitutions of the invariant cysteine residues Arg373 and Arg380 were made in the context of the Cdc7(ΔN)-Dbf4(MC) construct and tested in the using an *in vitro* kinase assay. Activity is normalized to the activity of Cdc7(ΔN)-Dbf4(MC). Error bars represent standard deviations determined from three replicates of a single experiment. A) Kinase assay using a peptide containing the primed Mcm2-S40 site with pre-phosphorylated S41. B) Kinase assay using a peptide containing the SE site Mcm2-S53.

4.5 Crystallisation and crystal screen optimisation of the ADP-BeF_3^- bound Cdc7-Dbf4 complex

4.5.1 Crystallisation of Cdc7(ΔN/2aq/3e)-Dbf4(MC)$^{ADP\text{-}BeF3}$

Crystal screening for Cdc7(ΔN/2aq/3e)-Dbf4(MC)$^{ADP\text{-}BeF3}$ was performed using a microseed matrix approach. The method involves adding a small amount of a seed stock to each condition of a sparse matrix screen. The seed stock is made by crushing crystals of the identical or a similar protein construct. Seeding often leads to rapid nucleation of protein crystals, allowing the second series of drops to have a sufficiently low precipitant concentration to equilibrate directly into the metastable zone enabling formation of crystals that may otherwise be impossible to obtain (see section 2.6.1). Microseed matrix screening has been used successfully in the crystallisation of a diverse range of proteins (Ireton and Stoddard, 2004, Obmolova et al., 2010, Till et al., 2013, D'Arcy et al., 2014, Kolek et al., 2016). The use of microseed matrix screening was particularly promising in this case as the seed stock could be made from crystals of an identical Cdc7-Dbf4 construct to that which was undergoing crystallisation screening. As the ligand was the only difference between the two complexes, the probability of the XL413-bound crystals nucleating the growth of the ADP-BeF_3^- bound kinase was deemed high.

Crystals of Cdc7(ΔN/2aq/3e)-Dbf4(MC)^{XL413} harvested from mother liquor were crashed by vortexing with a seed bead (Hampton Research) to create a seed stock. A mosquito robot was programmed to supplement each 400-nl sitting drop with 10 nl of seed stock. Cdc7(ΔN/2aq/3e)-Dbf4(MC)$^{ADP\text{-}BeF3}$ was also supplemented with various substrate peptides (2.5mM) in an attempt to make transition complexes with the kinase (Chapter 5). The obtained crystals were grown in the presence of the S53L peptide (Table 5-2), derived from the N-terminal region of Mcm2.
The crystals of Cdc7(ΔN/2aq/3e)-Dbf4(MC)$^{ADP\text{-}BeF3\text{-}}$ were obtained in a number of crystallisation conditions with acidic pH. Numerous purifications of the complex were used to optimise buffer pH, precipitant concentration, salt concentration, $MgCl_2$ and $ZnCl_2$ concentrations at 18°C. Further to this, additive and ionic liquid

screens were utilised and the concentrations of beneficial additives were also optimised against the precipitant concentration. The most readily optimised crystals grew in 21% PEG 1500, 20 mM $MgCl_2$,15% v/v acetonitrile, and 0.1 M DL-malic acid/MES/Tris (MMT) buffer pH 6.57. Crystals were grown in manual 2-µl hanging drops at 18°C and generally grew to a full size within 16 h. Microseeding was still required in manual optimisation plates, and was done using a microseeding fibre (Hampton Research).

4.5.2 Optimisation of cryo conditions for Cdc7(ΔN/2aq/3e)-Dbf4(MC)$^{ADP\text{-}BeF3}$

Crystals were harvested from wells and moved through a number of drops containing 75 mM NaCl, 20 mM $MgCl_2$, 1 mM ADP, 70 mM MMT buffer pH 6.57, 5 mM Tris-HCl, 15% acetonitrile, 1 mM BeF_3 and 0.5 mM S53L peptide with stepwise increases in PEG1500 (19-24% in 1% steps) and the cryo-protectant MPD (5-25% in 5% steps). An optimal freezing condition was achieved at 22% PEG1500 and 20% MPD.

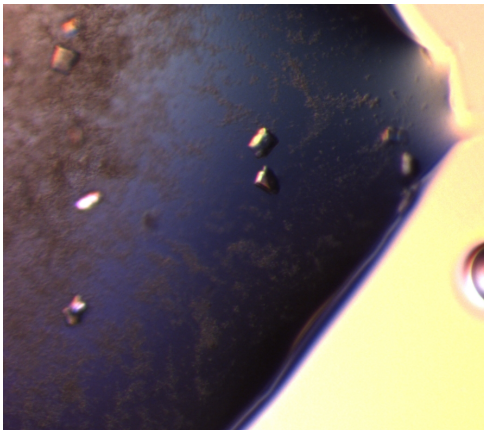


Figure 4-10 Crystals of Cdc7(ΔN/2aq/3e)-Dbf4(MC)$^{ADP\text{-}BeF3}$grown in the presence of ADP-BeF_3^-

Crystals were grown by vapour diffusion in 2ul hanging drops in 0.1 M MMT buffer pH 6.57, 21% PEG 1500, 15% v/v acetonitrile and 20 mM $MgCl_2$. Crystals were grown at 18°C and harvested after 48 h.

4.6 X-ray data collection and refinement of the Cdc7(ΔN/2aq/3e)-Dbf4(MC)$^{ADP\text{-}BeF3}$ structure

4.6.1 Data Collection

Data was collected at the macromolecular crystallography beamline I03 of the Diamond Light Source (Didcot, UK). I03 has a tuneable working wavelength range of 0.6-2.48 Å and is equipped with a Pilatus3-6M direct detector. All data was collected at the standard working wavelength of 0.976 Å (12.7 keV)
The optimised crystallisation and cryo-cooling conditions allowed for data to be obtained at a resolution of 1.79 Å (Figure 4-11).

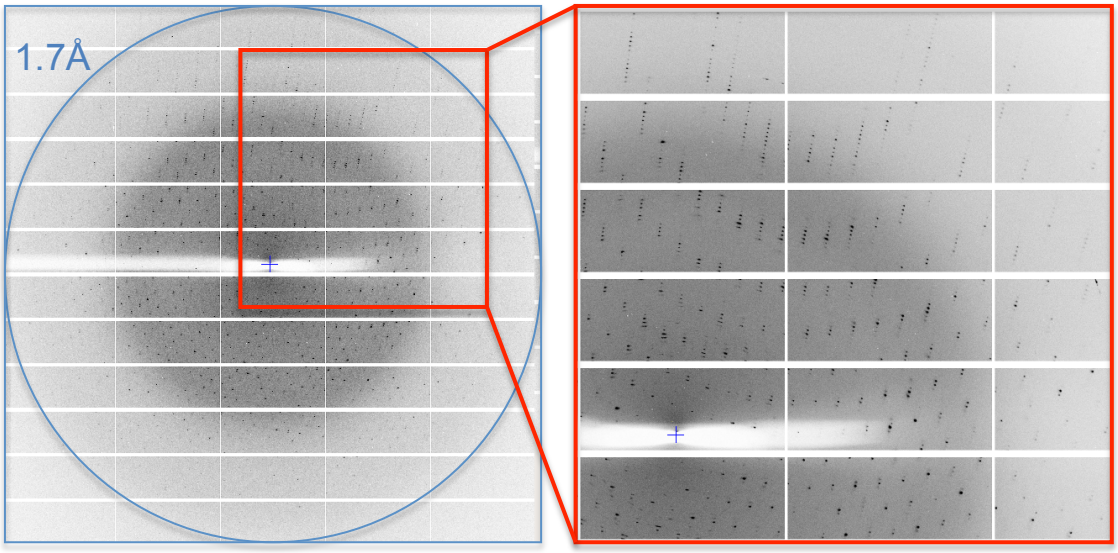


Figure 4-11 A sample X-ray diffraction image from a crystal of Cdc7(ΔN/2aq/3e)-Dbf4 (MC)$^{ADP\text{-}BeF3}$

A sample diffraction pattern is shown for crystals of Cdc7(ΔN/2aq/3e)-Dbf4 (MC) in complex with ADP-BeF_3^-, a full data set was collected at beamline I03 at the Diamond Light Source (DLS) to 1.7Å.

4.6.2 Data processing and structural refinement

Cdc7(ΔN/2aq/3e)-Dbf4(MC)$^{ADP\text{-}BeF3}$ also crystallised in the space group $P4_12_12$ with one molecule of the kinase present in the asymmetric unit. The collected X-ray diffraction data was processed using XDS (Kabsch, 2010b) and Scala (Evans, 2006b) via the Xia2 automatic data reduction pipeline (Winter, 2010b). The structures were solved by molecular replacement in Phaser (McCoy et al., 2007), and built using Phenix Autobuild (Adams et al., 2010a). The previous crystal structure of Cdc7(ΔN/2aq/3e)-Dbf4(MC)^{XL413} was used as a model. The structure was refined using phenix.refine (Adams et al., 2010a) and Refmac (Murshudov et al., 1997) with manual model building in Coot (Emsley and Cowtan, 2004a). The structure was refined to 1.79Å. X-ray data collection and refinement statistics are given in table 4-2.

4.7 Structure of (ΔN/2aq/3e)-Dbf4 (MC)$^{ADP\text{-}BeF3}$

4.7.1 Overall architecture of (ΔN/2aq/3e)-Dbf4 (MC)$^{ADP\text{-}BeF3}$

The overall architecture of (ΔN/2aq/3e)-Dbf4 (MC)$^{ADP\text{-}BeF3}$ is similar to that of (ΔN/2aq/3e)-Dbf4 (MC)^{XL413} aside from a small difference in orientation between the N and C lobes of the kinase. ADP-BeF_3^- is bound in the ATP binding pocket as expected alongside the two Mg atoms required for phosphoryl transfer. Although the crystals were obtained and optimized in the presence of an Mcm2-derived peptide, no density attributable to the peptide was not observed. This is not unexpected as kinase-substrate interactions are meant to be rapid and therefore the affinity of a kinase for a substrate peptide is often low, with interactions away from the active site likely contributing more to the affinity of a kinase for its substrate than those in the immediate vicinity of the active site.

4.7.2 Active site of Cdc7 (ΔN/2aq/3e)-Dbf4 (MC)$^{ADP\text{-}BeF3}$

The previous structure of Cdc7-Dbf4, solved with AMP-PNP, contained no density for the γ-phosphate, either due to flexibility of the moiety or due to the ability of Cdc7 to hydrolyse AMP-PNP. The structure consequently contained only one of the two Mg^{2+} ions required for the phosphor-transfer reaction (Hughes et al., 2012). In the current structure the BeF_3^- moiety occupies the position of the ATP γ-phosphate, flanked by the two necessary Mg atoms (Figure 4-12).

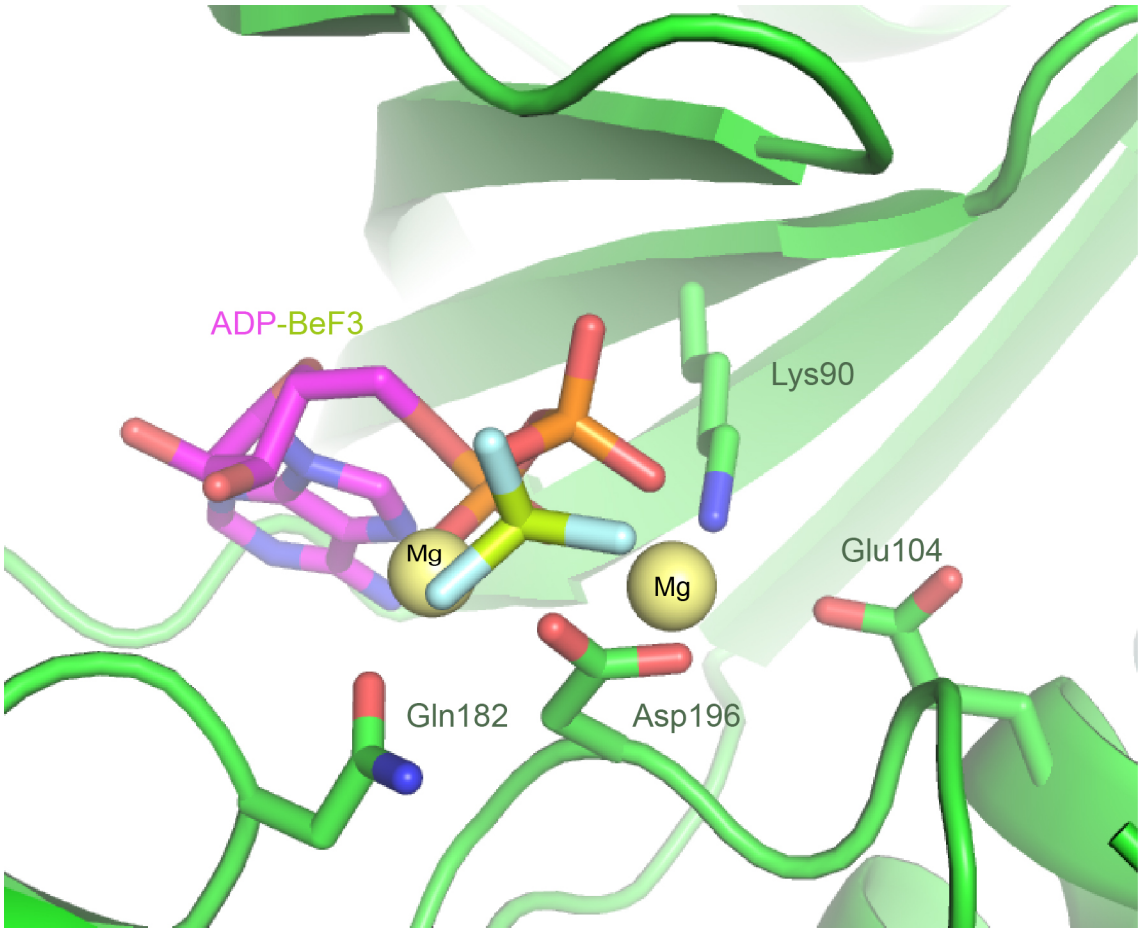


Figure 4-12 Active site of Cdc7-Dbf4 in complex with ADP-BeF_3^-

ADP-BeF_3^- bound in the active site and selected Cdc7 residues are and shown as sticks. Mg atoms are shown as yellow spheres.

4.8 Conclusions

In this chapter a method for obtaining diffraction quality crystals of a minimal Cdc7-Dbf4 construct with apparent WT levels of kinase activity has been described. The construct allowed determination of two near-atomic structures of the kinase in complex with XL413 and the ATP mimic ADP-BeF_3^-. Structural information gained in this chapter has shed new light on the features of Cdc7 critical for full activation of the kinase upon Dbf4 binding. A previously unseen Zn-binding domain in KI2 sandwiches Motif-M of Dbf4 between the C-lobe of Cdc7 and the new domain. This leads to full engagement of the kinase with its activating subunit. Importantly, this new interaction pins back and orders the activation loop of Cdc7. This newly ordered portion of KI2 appears to form a peptide-binding platform, onto which peptides from other kinase structures can be successfully docked. The structural importance of the KI2 Zn-binding domain correlated well with mutagenesis analysis in chapter 3, which showed the Cys residues coordinating the Zn atom are essential for *in vitro* kinase activity.

The structures also suggested a possible role for two Cdc7 residues Arg373 and Arg380 in substrate recognition. These residues, conserved across all known Cdc7 orthologs, seem to be ideally positioned to make favourable interactions with the P+1 residue of the substrate, which in the case of Cdc7 is occupied by an Asp, Glu, or a pre-phosphorylated Ser residue. However, without a substrate-bound structure of Cdc7-Dbf4 it is impossible to infer any further important features required for substrate binding. This poses further questions, which will be addressed in Chapter 5.

The first structure of Cdc7-Dbf4 bound to an ATP analogue is also presented. Previous structures of Cdc7-Dbf4 lacked the γ-phosphate of the analogue possibly due to hydrolysis during crystallization. In the newly obtained structure, BeF_3^- mimics the ATP γ-phosphate and is present along with a pair of Mg atoms. These Mg atoms overlay perfectly with those in the transition state crystal structure of CDK2-CyclinA obtained by Bao et al. (2011). As the most abundant divalent metal ion in the cell, Mg^{2+} presents the most obvious choice for a metal co-factor for kinase activity. However, kinases are able to utilise a number of divalent metal ions

for activity, which can be used as a method of fine-tuning kinase activity. Structures of the cyclic AMP (cAMP) dependent PKA have been refined with product complexes utilising a number of cofactors including Mg^{2+}, Ca^{2+}, Sr^{2+} and Ba^{2+} (Gerlits et al., 2013). The ability of PKA to utilise Ca^{2+} for phosphoryl transfer may have physiological relevance. This kinase is localised near calcium channels in the cell, suggesting PKA as a possible link between cAMP and calcium signalling. Calcium has been shown to decrease the affinity of PKA for pseudosubstrate inhibitors and, while it is able perform its catalytic function, it inhibits the release of the phosphorylated substrate, leading to reduced turnover rates (Knape et al., 2015). In this respect, choice of metal ion could be used for kinase regulation in response to different cellular signals.

The two metal mechanism of catalysis is also utilised by CDK2, which requires two Mg^{2+} ions to function. Binding of a second Mg^{2+} is proposed to create a more rigid and enclosed active site to better shield the γ-phosphate from solvent while stabilising the build-up of negative charge. This allows for faster catalysis to occur (Bao et al., 2011) However, the presence of two such ions actually limits the rate of turnover as the second Mg^{2+} ion stabilises the produced ADP in the binding pocket and reduces its dissociation rate. This product release is thought to be the rate-limiting step for CDK2 (Jacobsen et al., 2012), revealing a fine balance between the kinetics of both substrate binding and phosphoryl transfer to maintain sufficient levels of kinase activity.

.

Chapter 5. Results 3 – Crystallisation of a Cdc7-Dbf4 heterodimer in complex with an *ATPγS-conjugated peptide*

5.1 Aims

Structures determined in Chapter 4 revealed the active site of Cdc7 in a fully open state and a potential substrate-binding platform. The next aim of the project was to elucidate the structural features of the kinase required for substrate recognition and binding. The optimised construct Cdc7(ΔN/2aq/3e)-Dbf4(MC) was maintained and efforts were made to identify suitable peptide substrates for co-crystallisation. Various methods of forming transition complexes between the kinase and peptide substrates were attempted to increase the affinity of the interaction to aid in co-crystallisation. A range of different target sites were used to try and ascertain the possible variations in binding requirements between phosphorylation-primed sites and SD/E target sites.

In this chapter a crystal structure of Cdc7(ΔN/2aq/3e)-Dbf4(MC) bound to an Mcm2-derived peptide is presented. The structure confirms the hypothesised importance of Arg373 and Arg380 in substrate recognition and reveals a further binding pocket in Cdc7 that accommodates the P+4 residue.

5.2 Identification and optimisation of Cdc7 substrate peptides

The Mcm2-7 complex, the heterohexameric helicase responsible for DNA unwinding during replication, is believed to be the main target of Cdc7-Dbf4. The phosphorylation of MCM2-7 is thought to trigger conformational changes in the helicase en route to origin firing in S-phase (Lei et al., 1997, Cho et al., 2006, Masai et al., 2006, Francis et al., 2009, Sheu and Stillman, 2010). The majority of phosphorylation sites map to the N-terminal tails of the Mcm2-7 subunits (Randell

et al., 2010). In yeast, the predominant target of Cdc7-Dbf4 appears to be Mcm4, the phosphorylation of which appears to relieve an auto inhibitory activity contained in the subunit, thereby facilitating Mcm2-7 activation (Masai et al., 2006, Sheu and Stillman, 2006, Randell et al., 2010, Sheu and Stillman, 2010). However, in human cells, the predominant target appears to be Mcm2, although Mcm2, 4 and 6 were all implicated as targets in both systems (Montagnoli et al., 2006, Tsuji et al., 2006, Charych et al., 2008, Chuang et al., 2009). To confirm known and identify potential novel Cdc7-dependent phosphorylation sites in the human replicative helicase, we used a library of synthetic peptides spanning the N-terminal 240 amino acids of MCM2, 4 and 6. The libraries were synthesised by the Peptide Synthesis Facility (London Research Institute, London) immobilised on a membrane support.

5.2.1 Peptide arrays of the N-terminal tails of Mcm2, 4 and 6 highlight numerous potential Cdc7 phosphorylation sites

Peptide arrays were used for radioactive kinase assays utilising a recombinant preparation of Cdc7(ΔN)-Dbf4(MC) kinase. A detailed description of the peptide array kinase assay is given in 2.4.5.1. In brief, the membrane was hydrated and equilibrated with kinase buffer in the presence of blocking agent to prevent non-specific binding of the kinase. Following incubation in the presence of [γ-^{32}P]ATP and recombinant kinase, the membrane was washed, and incorporated radioactivity was quantified by phosphor imaging (Figure 5-1). A number of Cdc7 phosphorylation sites were identified as part of the array (Table 5-1), and some of these were synthesised in various forms for co-crystallisation trials (Table 5-2).

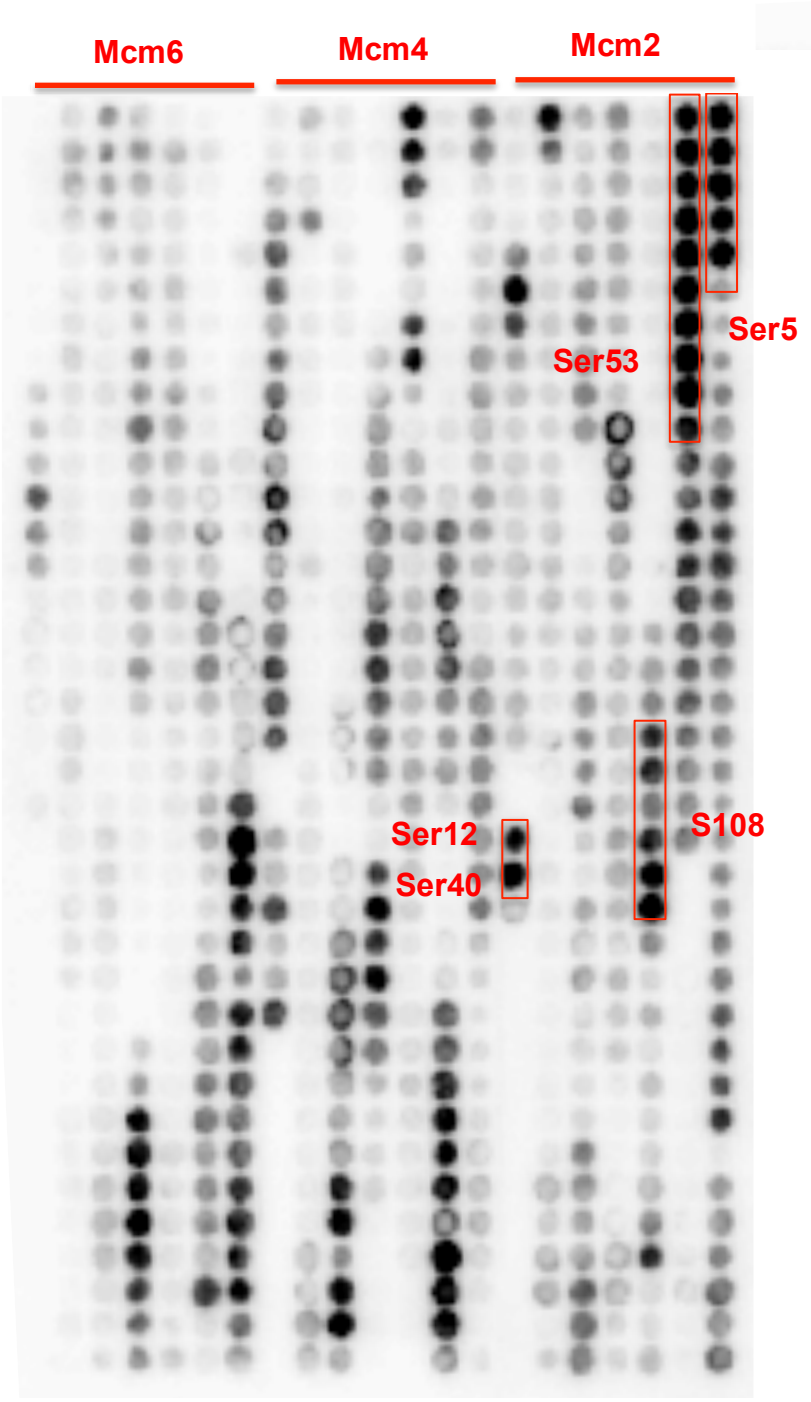


Figure 5-1 A peptide array kinase assay revealed a number of potential target sites for Cdc7-Dbf4

Each dot represents a 20-mer peptide from N-terminal portions of human Mcm2, 4 or 6 shifting one amino acid residue along the protein sequence with each dot. Peptides containing potential CDK2-primed Cdc7 phosphorylation sites (Ser-Ser-Pro) sites were synthetized with phosphoserine residues to mimic priming by CDK2. Rows of dark dots represent single phosphorylation sites. Example sites are indicated for Mcm2 in red boxes. A comprehensive list of sites is given in table 5-1.

Table 5-1 Table of Cdc7 phosphorylation sites identified using peptide arrays

Mcm2	Mcm4	Mcm6
S5 - SE	S31- SSp	S40 - SD
S12 – SSp	S70- SS	S76- ST
S40 – SSp	S87- SSp	S153- SG
S53 - SE	S131- SD	
S108 - SQ	S142- SE	
S139 - SD	T189- TE	
S187 - SM	S228 -SY	
S229 - SL	T236- TF	

* The target and P+1 residues are shown for each site (Sp = phosphor-serine). Sites highlighted in red do not follow the expected consensus sequence of Cdc7.

Table 5-2 Table of peptides used for crystallography trials

Peptide Name	Target Site	Residues	Modification
Mcm2(34-46)	S40	34-46	S41-P
Mcm2(36-44)	S40	36-44	S41-P
Mcm2(37-43)	S40	37-43	S41-P
Mcm2(33-47)	S40	33-47	S41-P
Mcm2(49-57)	S53	49-57	n/a
Mcm2(50-56)	S53	50-56	n/a
S53(19)	S53	44-62	n/a
S5(9)	S5	1-9	n/a
S40-ATPγS(9)	S40	36-44	ATPγS/S41-P
S40-ATPγS(15)	S40	33-47	ATPγS/S41-P
S53-ATPγS(9)	S53	49-57	ATPγS
S53-ATPγS(15)	S53	46-60	ATPγS
S5-ATPγS(9)	S5	1-9	ATPγS
S5-ATPγS(15)	S5	1-15	ATPγS

5.3 Production of complexes of peptides with Cdc7 (ΔN/2aq/3e)-Dbf4 (MC)

5.3.1 Crystal screening of transition complexes using peptides with ADP-AlF_4^- and ADP-BeF_3^-

Crystal screening for substrate-bound Cdc7 (ΔN/2aq/3e)-Dbf4 (MC) was carried out with microseeding using Cdc7 (ΔN/2aq/3e)-Dbf4 (MC)^{XL413} crystals as a seed stock. Complexes of Cdc7-Dbf4 were made in multiple combinations with different peptides and equimolar concentrations of ADP-BeF_3^- or AlF_4^-. Each transition complex was also screened in the presence or absence of $MgCl_2$ (2.5-20 mM).

A number of conditions yielded good quality crystals, allowing for acquisition of X-ray diffraction data on synchrotron sources. However, upon data processing and phasing by molecular replacement, electron density that could be attributed to the bound peptide was not observed. Following multiple failed attempts, it was hypothesized that the peptides did not bind the kinase tightly enough, resulting in low occupancy. The length of the peptide seemed to have little effect on crystal formation, and increasing the peptide length to potentially enhance affinity did not yield substrate bound crystals. Cdc7-Dbf4 displays a high affinity for ATP, displaying a Km of 2-4 μM (Masai et al., 2000). We reasoned that covalently attaching a substrate peptide to the nucleotide would improve the chances of visualising a bound substrate. Indeed, such an approach proved successful in the crystallisation of the insulin receptor protein Tyr kinase (IRK) (Hines et al., 2005). This structure was useful in establishing important features in substrate binding and possible considerations for designing inhibitors. Substrates for both Ser/Thr and Tyr kinases have been produced with considerably improved affinities through tethering of nucleotide analogues and ATP-competitive inhibitors to kinase substrate peptides. In one of the early attempts, the affinity of one such bisubstrate ligand of PKA was shown to be 67-fold greater than of the ATP analogue alone (Ricouart et al., 1991). More recently, a similar approach was used to tether an ATP competitive small molecule to a substrate peptide of PKC, leading to improved affinity and selectivity for specific PKC isoforms relative to the inhibitor or peptide alone (van Ameijde et al., 2010). Success was also achieved through covalently

linking of inhibitors to peptides designed for binding of substrate docking sites away from the active site. When applied to the MAPK c-Jun N-terminal kinase 1 (JNK1) this approach created a bisubstrate inhibitor for the kinase with a 20,000-fold higher affinity than the small molecule alone (Stebbins et al., 2011).

5.3.2 Design and production of ATPγS conjugated peptides.

Modified peptides were synthesised by the LRI Peptide Synthesis Facility. Full details for the production of the ATPγS conjugated peptides can be found in section 2.4.5. In brief, peptides were produced by solid phase peptide synthesis using standard Fmoc chemistry with an aminoalanine in place of the target Ser residue. The non-standard residue was bromoacetylated post-synthetically by incubating the peptide with bromoacetic acid. The ATPγS conjugates were then made by reacting the sulfhydryl group of ATPγS to the bromoacetyl moiety. The procedure produced a covalent bond linking the nucleotide analogue and the peptide, creating a bi-substrate ligand (Figure 5-2).

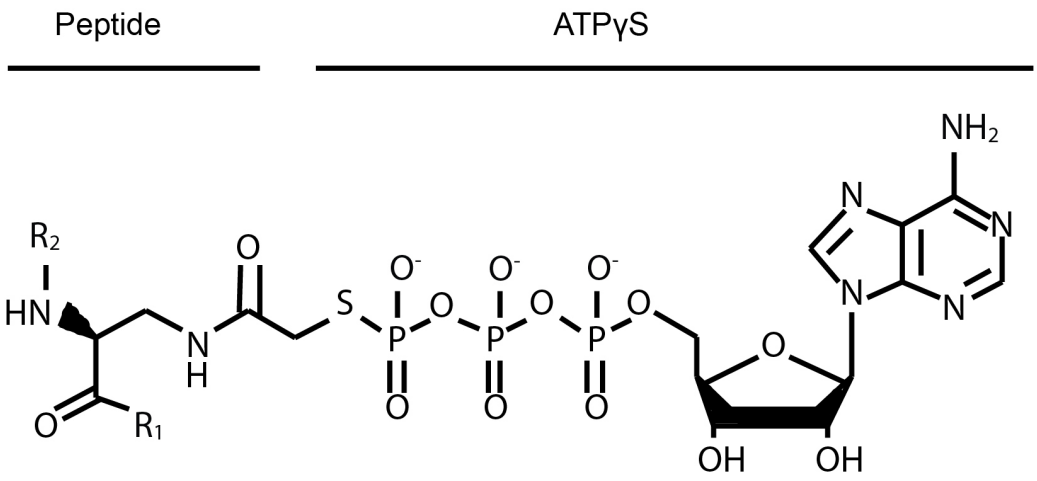


Figure 5-2 Generalised structure of a bi-substrate inhibitor peptide
R_1 and R_2 represent N- and C- portions of the substrate peptide**.**

5.4 Crystal screens optimisation and freezing of (ΔN/2aq/3e)-Dbf4 (MC) $^{ATPγS\text{-}Mcm2\text{-}S40(15)}$

Cdc7 (ΔN/2aq/3e)-Dbf4 (MC) supplemented with 1-2.5 mM nucleotide-peptide conjugate was subjected to microseed matrix screening using crystals of Cdc7(ΔN/2aq/3e)-Dbf4 (MC)^{XL413} as a seed stock. A number of hits were obtained in the initial screens using the ATPγS-Mcm2-S40(15) peptide. The most promising condition contained 25% PEG 1500 in 0.1 M propionic acid/cacodylate/BisTris propane (PCTP) buffer, pH 7. Manually setup grids and streak microseeding was used to optimise the pH of PCTP buffer against the precipitant concentration. Additive screens were employed at later stages, and the additive concentration (MPD) was optimised against the precipitant concentration. The final optimised condition contained 18% PEG 1500, 8% MPD in 0.1 M PCTP buffer, pH 6.5. Crystals were grown by vapour diffusion in hanging drops at 18°C and typically appeared within 16 h (Figure 5-3). Crystals were harvested within 16-24 h and snap-frozen in liquid nitrogen using the cryo-conditions described in 5.4.1.

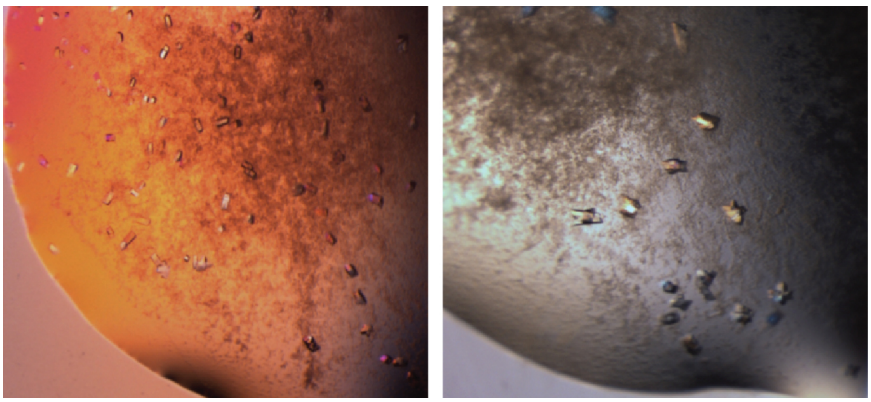


Figure 5-3 Optimised crystals of Cdc7 (ΔN/2aq/3e)-Dbf4 (MC) grown in the presence of ATPγS-Mcm2-S40(15).

5.4.1 Cryo-condition optimisation for data collection

Crystals, harvested from mother liquor were incubated in 5-μl drops containing 1 mM ATPγS-Mcm2-S40(15), 75 mM NaCl and 80 mM PCTP pH6.5 with stepwise increases in PEG 1500 (17-21% in 1% steps) and the cryo-protectant MPD (10-25% in 5% steps). An optimal freezing condition was achieved in the presence of 19% PEG1500 and 20% MPD, whereby no crystalline ice was observed.

5.5 Data collection and structure refinement of Cdc7 (ΔN/2aq/3e)-Dbf4(MC) $^{ATPγS\text{-}Mcm2\text{-}S40(15)}$

5.5.1 Data collection for Cdc7 (ΔN/2aq/3e)-Dbf4 (MC) $^{ATγS\text{-}Mcm2\text{-}S40(15)}$

The optimised crystallisation and cryo-cooling conditions allowed for data to be obtained at a resolution of 1.67 Å on the beamline I03 of the Diamond Light Source (Didcot, UK) (Figure 5-4). I03 is macromolecular crystallography beamline with a tuneable working wavelength range of 0.6-2.48 Å, equipped with a Pilatus3-6M direct detector. All data were collected at the standard working wavelength of 0.976 Å (12.7 keV). Crystals from similar conditions diffracted in the range of 1.8-2.5 Å. The dataset at 1.67 Å was considered to be of the best quality and was subsequently processed and used to solve the structure of Cdc7 (ΔN/2aq/3e) - Dbf4 (MC) $^{ATPγS\text{-}Mcm2\text{-}S40(15)}$.

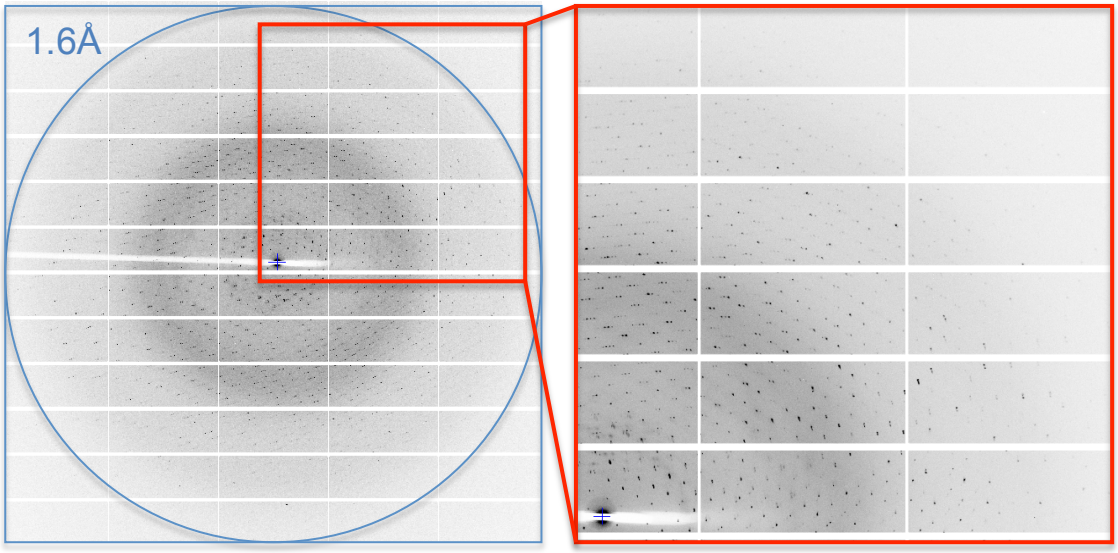


Figure 5-4 A sample X-ray diffraction pattern collected from a crystal of (ΔN/2aq/3e)-Dbf4(MC)$^{ATPYS-Mcm2-S40(15)}$

5.5.2 Structure refinement of Cdc7 (ΔN/2aq/3e)-Dbf4 (MC)$^{ATPγS-Mcm2-S40(15)}$

The collected X-ray diffraction data was processed using XDS (Kabsch, 2010b) and Scala (Evans, 2006b) via the Xia2 automatic data reduction pipeline (Winter, 2010b). The structure was solved by molecular replacement in Phaser (McCoy et al., 2007), and built using Phenix Autobuild (Adams et al., 2010a). The previous crystal structure of Cdc7 (ΔN/2aq/3e)-Dbf4 (MC)$^{ADP-BeF3}$ was used as a molecular replacement model. The structure was refined using phenix.refine. (Adams et al., 2010a) and Refmac (Murshudov et al., 1997) with manual model building in Coot (Emsley and Cowtan, 2004a). The structure was refined to 1.67 Å. X-ray data collection and refinement statistics are given in Table 5-3.

Table 5-3 Data collection and refinement statistics for Cdc7(ΔN/2aq/3e)-Dbf4(MC) bound to ATPγS-Mcm2-S40(15)

	ATPγS-Mcm2-S40(15)
Data collection *	
Beamline	DLS I03
Space group	P 41 21 2
Unit cell parameters	
a, b, c (Å)	62.28, 62.28, 234.79
α, β, γ (^{o})	90, 90, 90
Wavelength (Å)	0.9796
Resolution range (Å)	58.7 -1.67 (1.71 -1.67)
Average multiplicity	12.1 (8.5)
Completeness (%)	100.0 (99.8)
CC1/2	0.999 (0.81)
R_{merge}	0.90 (0.724)
R_{meas}	0.097 (0.821)
</>/<sig(*I*)>	19.4 (3.3)
Wilson <B> (Å^{2})	17.23
Solvent content (%)	42.5
Refinement	
Resolution	**58.7 -1.67**
Reflections (total/free)	665576 (55037)
R_{work}/R_{free}	0.1602/0.1847
No. Atoms	4043
Protein	3653
Ligand/ion	36
Water	354
Average B-factors	23.84
Protein	22.97
Ligand/ion	30.13
Water	32.21
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	1.08
Ramachandran (%)	
Favoured	97
Outliers	2.6
Number of TLS groups	6

*Data for the highest resolution shell is given in parentheses.

5.6 Structure of (ΔN/2aq/3e)-Dbf4 (MC) $^{ATPγS\text{-}Mcm2\text{-}S40(15)}$

5.6.1 Substrate binds in the peptide-binding platform created by the Zn binding domain in Cdc7

Cdc7 (ΔN/2aq/3e)-Dbf4 (MC) $^{ATPγS\text{-}Mcm2\text{-}S40(15)}$ crystallised in the space group $P4_12_12$ with one molecule of the kinase-substrate complex in asymmetric unit. The overall structure of the Cdc7-Dbf4 construct is very similar to that observed in complex with XL413, with r.m.s. deviations for common Cα atom positions of 0.345 Å (Figure 5-5). The initial difference *Fo-Fc* map revealed a pronounced additional density, which could be unambiguously interpreted as the peptide moiety of the chimeric ligand (Figure 5-6). In total, 9 residues of the peptide (corresponding to Mcm2 residues 38-46), including the aminoalanine mimicking the target Ser40 could be built (Figure 5-6A).

While the peptide and nucleotide are both bound in the active site as expected, it can be seen that hydrolysis of the bi-substrate inhibitor has occurred during the crystallisation process (Figure 5-6B). The hydrolysis left ADP in the active site with the peptide containing a sulfhydryl group which both interact with a metal atom in the active site. The metal atom was interpreted as Zn, which was present in the protein sample.

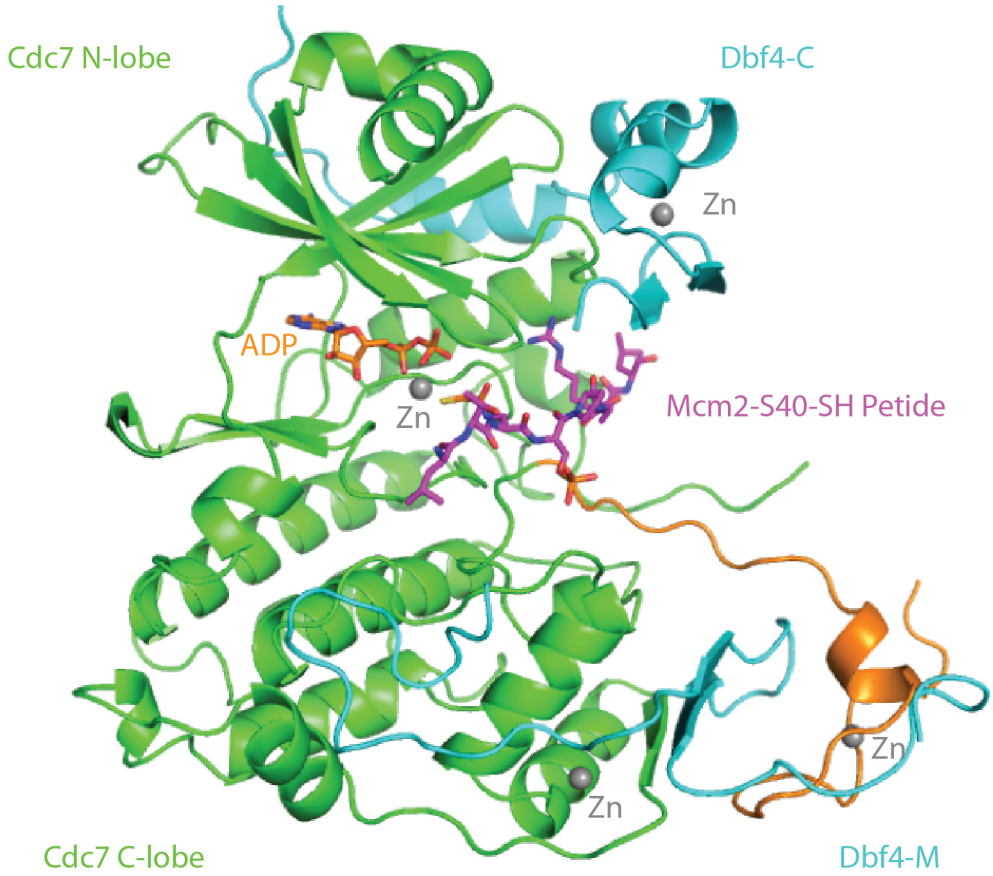


Figure 5-5 Overall structure of Cdc7-Dbf4 bound to ATPγS-Mcm2-S40(15)

ADP and the Mcm2-derived peptide are shown as sticks in orange and magenta respectively. All non-carbon atoms are shown by standard colouration. Zinc atoms are shown as grey spheres.

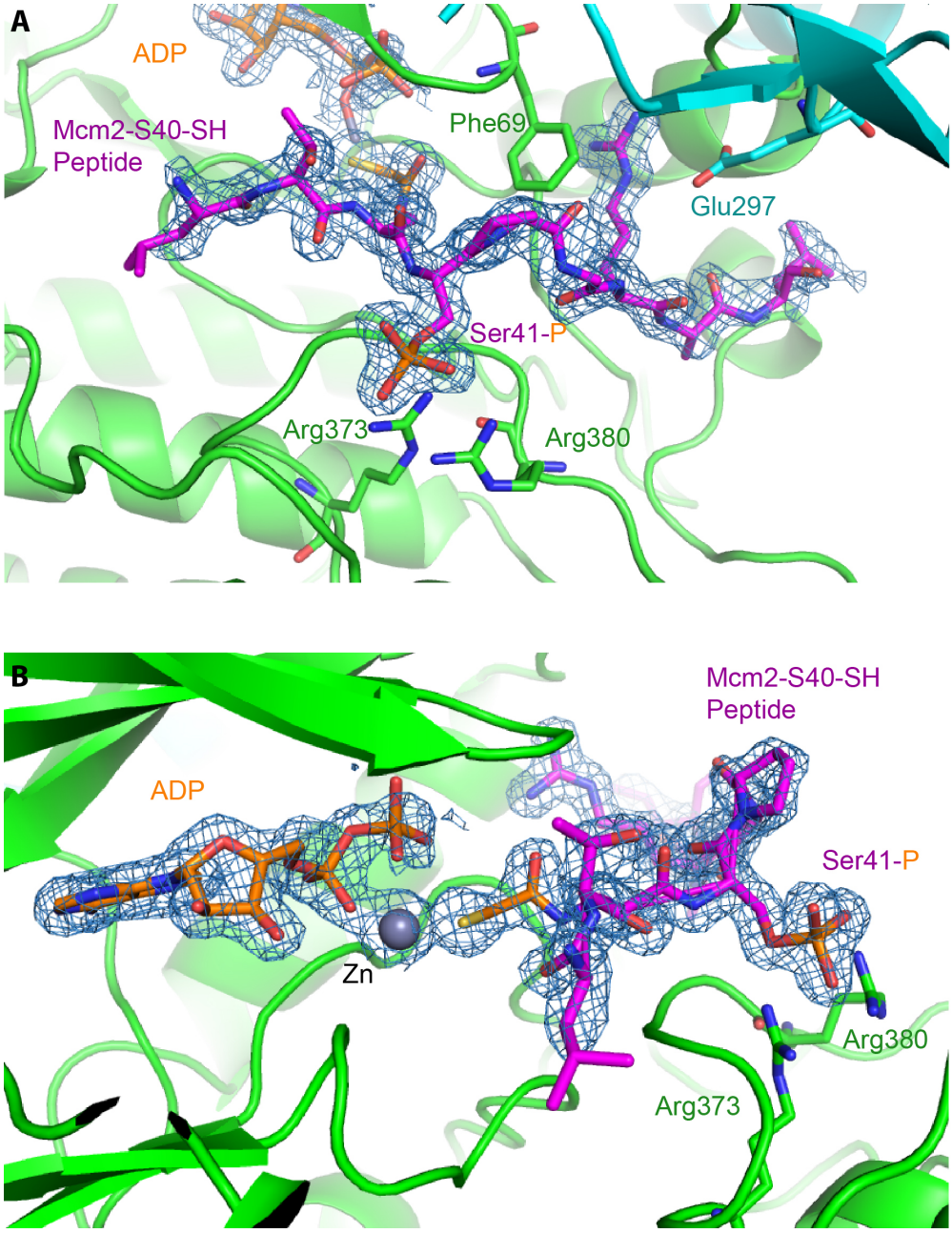


Figure 5-6 Electron density reveals successful substrate binding

A) The Mcm2-derived peptide binds in the cleft between the N- and C-lobe of Cdc7. B) The ATPγS-conjugated peptide appears to be hydrolysed during crystallisation. The ligand was therefore modelled as ADP and an SH-conjugated peptide held together by interaction with a metal atom (modelled as Zn), as suggested by the density. Residues discussed in the text are shown as sticks. Weighted *2Fo-Fc* electron density of the bound nucleotide and the peptide is shown as blue chicken wire, contoured at 1.0σ.

5.6.2 Arg373 and Arg380 interact with the pre-phosphorylated serine of the peptide required for substrate specificity

As hypothesized based on modelling in Chapter 4 (Figure 4-8), Arg373 and Arg380 make favourable interactions with the phosphate group of the P+1 residue of the peptide (Figure 5-7) This pre-phosphorylated or negatively charged P+1 residue constitutes the entire consensus sequence for Cdc7 phosphorylation sites (Cho et al., 2006, Montagnoli et al., 2006). The importance of this interaction in the structure is also in excellent agreement with mutagenesis analysis (Figure 4-9).

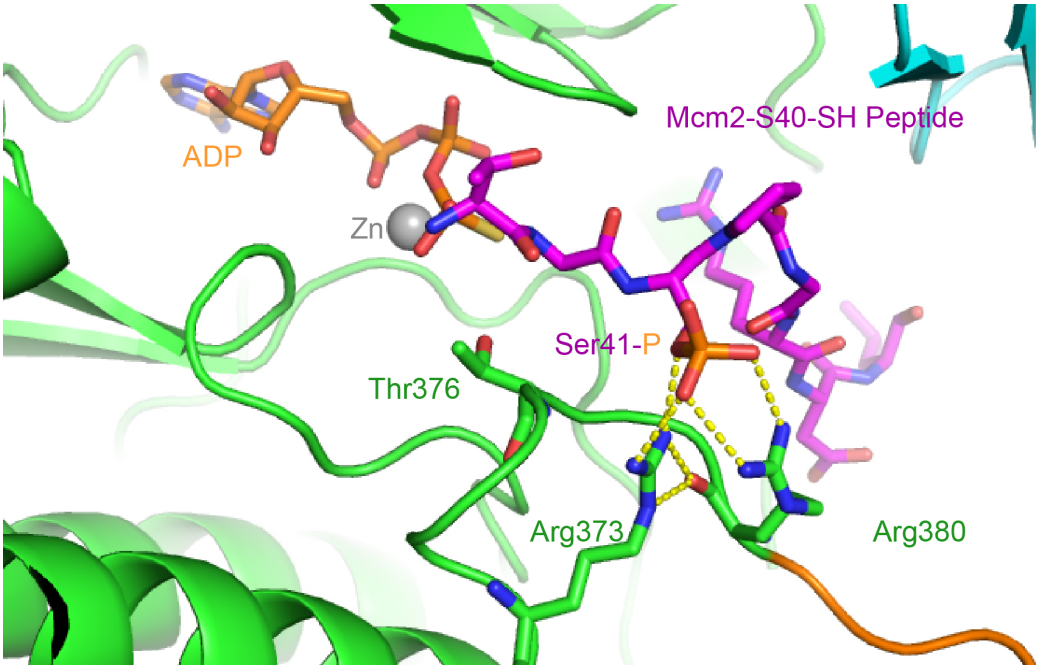


Figure 5-7 The structure confirms the recognition of the P+1 residue by Cdc7 Arg373 and Arg380 and suggests a possible role for Thr376

Residues discussed in the text are shown as sticks and indicated. Yellow dashes represent hydrogen bonds and salt bridge interactions involving Arg residues of Cdc7 and the phosphate group of the peptide.

5.6.3 Cdc7 contains a binding pocket for the P+4 residue of the target suggesting a co-evolution with CDK2

The P+4 position in the S40(15) model substrate is occupied by an Arg residue, corresponding to Arg44 of Mcm2. In the structure, the guanidinium side chain of Arg44 stacks against Cdc7 Phe69 and forms hydrogen bonds with carbonyl groups of Cdc7 Ser96 and Gly198 and a salt bridge with Dbf4 Glu297 (Figure 5-8). These Cdc7 and Dbf4 residues form a favourable pocket for a positively charged side chain at the substrate P+4 position. Intriguingly, the target site consensus of CDK2, the kinase proposed to prime Ser40 for recognition by Cdc7 by pre-phosphorylating Ser41, features an Arg or a Lys residue at P+3 position (Stevenson-Lindert et al., 2003). Therefore, it appears that Cdc7 is specifically adapted to act upon sites primed by CDK2, suggesting a co-evolution of the S-phase kinases (Nougarede et al., 2000). It is however worth noting that not all Cdc7 sites are primed sites, and a range of residues can be present at the P+4 position of any site. This was confirmed by *in vitro* kinase assays performed on peptides with various substitutions at the P+4 residue, in which mutation of the arginine residue has minimal effects on activity in many cases (Figure 5-9). The decrease in activity seen for some mutants is also significantly less than the decrease observed upon removal of the primed Ser41, highlighting how the P+1 residue is the essential requirement for substrate specificity and full activity. It is however important to consider that such activity assays as a method of inferring affinity does not dissect the relative on and off rates of the kinase-peptide interaction. It is possible that peptides that appear to have less affinity actually have a higher affinity for the kinase and are therefore released more slowly leading to slower turnover. To assess the affinity of the various mutant peptides for the kinase, biophysical studies such as ITC or SPR would be required.

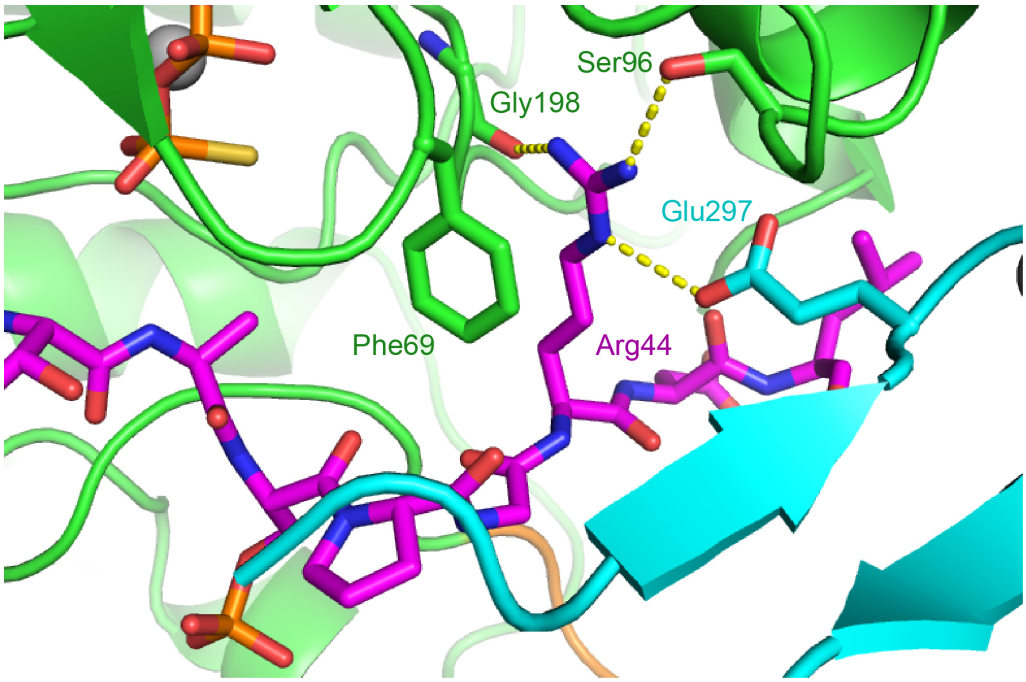


Figure 5-8 An arginine binding pocket allows for favourable interactions between the P+4 residue of the peptide and Cdc7-Dbf4.

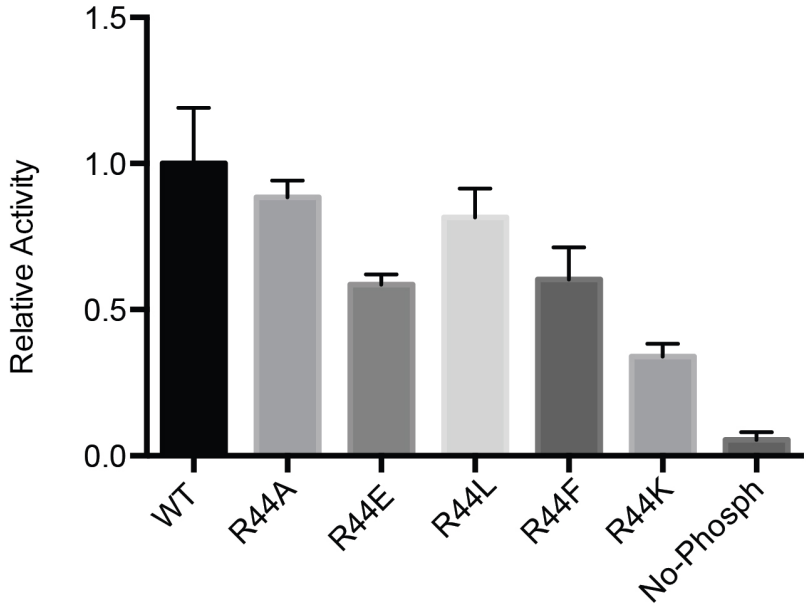


Figure 5-9 Arg44 of Mcm2 is not essential for phosphorylation of Ser41 by Cdc7-Dbf4 *in vitro*

Peptides with indicated substitutions of the P+4 residue were used as substrates for phosphorylation by Cdc7(ΔN)-Dbf4(MC) *in vitro.* The activity relative to that observed with WT Mcm2 peptide is shown in form of a bar chart. Error bars represent standard deviations calculated from a single experiment performed in triplicate. No-Phosph peptide contained Arg44 but no phosphorylation onf Ser41.

5.6.4 A possible role for Thr376 in kinase regulation

Activities of CDKs and numerous other kinases are regulated by phosphorylation of their activation loops by priming kinases leading to conformational changes required for kinase activation (Adams, 2003). In the case of CDK2-Cyclin A, phosphorylation of Thr160 allows the residue to contribute to a hydrogen bond network that facilitates interactions between the kinase and the activating cyclin. Phosphorylation moreover leads to a significant movement of the Thr residue, which opens the active site for substrate binding (Russo et al., 1996). The mitiotic kinase Aurora A also has a conserved threonine which is phosphorylated as part of its activation (Thr288) which forms important interactions with the αC helix upon binding of the activator TPX2. Binding of TPX2 induces a 10-Å movement of the phospho-Thr residue in order for the active conformation to be formed. In the absence of TPX2, Thr288 is solvent-exposed and available for phosphorylation.

Binding of TPX2 appears to enhance auto-phosphorylation and protect the phospho-Thr from phosphatases. This is in contrast to CDK2 which requires cyclin binding to facilitate Thr160 phosphorylation (Bayliss et al., 2003). Phosphorylation also plays an important role in the activation of Tyr kinases, with all MAP kinases requiring phosphorylation at two sites for full activation (Canagarajah et al., 1997, Bellon et al., 1999). In contrast, the lymphocyte tyrosine kinase (Lck) kinase requires a single phosphorylation at Tyr394 for activation, while a second phosphorylation event at Tyr505 inhibits the kinase activity (Yamaguchi and Hendrickson, 1996).

Masai and colleagues suggested that Cdc7 may similarly depend on phosphorylation of its activation loop and implicated a number of kinases including CDK2 as the possible priming kinase with Cdc7 Thr376 as the target (Masai et al., 2000). Mutagenesis confirmed that Thr376 is critical to Cdc7 activity, and it was shown to be a substrate for phosphorylation by a number of kinases *in vitro*. However, superposition of Cdc7 and phosphorylation-activated kinase structures revealed that Cdc7 Thr376 is unlikely to function as the activating Thr residue (Hughes et al., 2012), suggesting an alternative mode of regulation. Due to its proximity both to the γ-phosphate of the nucleotide and the negatively charged P+1 residue of the substrate, phosphorylation of Thr376 is likely to inhibit peptide binding and subsequent activity (Figure 5-7). It is therefore possible that phosphorylation of Thr376 acts as an off-switch for the kinase. This idea is supported by recent data from Lee and colleagues, who showed that phosphorylated Cdc7 has limited affinity for replication origins, and de-phosphorylation of a number of sites by PP1 may be required for DNA replication. They furthermore suggested that phosphorylation of Cdc7 by mitotic CDK1 serves to prevent DNA re-replication in mitosis (Knockleby et al., 2016).

Inhibitory phosphorylation of kinases is more commonly associated with the P-loop in the active site, with phosphorylation of residues in this glycine-rich region preventing proper alignment of ATP in the active site (Morgan, 1995). However, there are examples of inhibitory phosphorylation in the activation loop of kinases. The activity of CDK9 has been shown to be reliant on the dephosphorylation of Ser175 by PP1. Presence of a phosphate at this residue has been shown to be

inhibitory for the kinases ability to phosphorylate RNAPolII and promote HIV-1 transcription. This was further confirmed by a total loss of *in vitro* activity upon phosphomimetic mutation of this residue. Conversely, an alanine mutation of Ser175 did not lead to a reduction in CDK9 activity (Ammosova et al., 2011). A recent interesting example of inhibitory phosphorylation of the activation loop is ERK1. An initial phosphorylation event is required at the conserved TEY for activation of ERK1. However, the increase in its catalytic activity leads to steady auto-phosphorylation of the kinase at Thr207, which reduces its activity towards other substrates. This serves as a mechanism to gradually switch off kinase activity after the initial activation event (Lai and Pelech, 2016). This appears to be a mechanism of regulation that is conserved across a number of kinases, with the checkpoint kinase Chk2 and the centrosomal kinase Nek2 both containing activating phosphorylation sites and inhibitory auto-phosphorylation sites in the activation loop (Schwarz et al., 2003, Rellos et al., 2007). This kind of regulation is also seen with microtubule-associated protein (MAP)-microtubule affinity regulating kinase (MARK) which contains both an activation phosphorylation site in its activation loop (Thr208) alongside an inhibitory site (Ser212) (Timm et al., 2003). However, in the case of MARK, the inhibitory phosphorylation is executed by glycogen synthase kinase 3β (GSK3 β) (Timm et al., 2008). Whether such a method of regulation is relevant in the case of Cdc7 is not currently clear but it could offer an intriguing possibility for fine-tuning the activity of Cdc7 as it moves through the cell cycle.

5.7 Conclusions

This chapter describes a method for obtaining crystals of a peptide-bound form of Cdc7-Dbf4. Successful co-crystallization of a model target peptide was achieved through the use of a chimeric bi-substrate ligand, in which the nucleotide analogue and the peptide were covalently linked to increase the affinity of the initial binding to the kinase. Crystals of Cdc7-Dbf4 were obtained in the presence of a number of different peptides, designed based on previously known and newly identified Cdc7 target sites. However, most of the crystallization conditions did not reveal electron

density for the peptide. A single structure was obtained of the previously crystallised Cdc7-Dbf4 construct (Chapter 4) in the presence of an Mcm2-derived 15-mer peptide conjugated to ATPγS at the Mcm2 Ser40 Cdc7 target site with the priming phosphorylation at the P+1 site. The structure revealed a number of important features for substrate binding and kinase activation despite apparent hydrolysis of the peptide-ATPγS linkage. The hydrolysis was not entirely unexpected, as the chimeric molecule spontaneously hydrolyses over time. This problem was more pronounced with more acidic peptides, and the hydrolysis could have been increased by the slightly acidic pH required for crystallisation.

The newly obtained structure confirms the importance of Cdc7 Arg373 and Arg380 for the recognition of the negatively charged P+1 residue in Cdc7 target site consensus. Crucially, Arg373 and Arg380 are conserved across all known Cdc7 orthologs, and the P+1 residue is the entirety of the Cdc7 consensus sequence. The importance of the P+1 substrate residue is further shown by *in vitro* kinase assays in which a substantial loss of activity in the absence of Ser41 phosphorylation is observed. Further to this, the unexpected interaction involving the side chain of the P+4 residue of Mcm2 (Arg44), which becomes buried in a pocket at the Cdc7-Dbf4 interface. This Arg residue is also a part of the consensus of CDK2, which prefers an Arg or a Lys residue at the corresponding P+3 position. The structure therefore suggests that the pocket has evolved to facilitate Cdc7-Dbf4 to act in tandem with CDK2 during DNA replication initiation. However, *in vitro* kinase assays using peptides containing mutants of the P+4 residue revealed limited importance of an Arg at this position in the kinase, with mutation of the P+1 residue proving to be much more detrimental to kinase activity. While the interactions are well-defined in the structure, it is not unexpected that the pocket would able to accommodate a number of different residues as the P+4 residue is not an arginine in many of the S(D/E) sites that Cdc7 also phosphorylates.

Another important feature of the structure is the proximity of Thr376 relative to the γ-phosphate of the nucleotide and the negatively charged P+1 residue of a bound substrate. The position of Thr376 is incompatible with the hypothesis of it being a phosphorylation-priming site for Cdc7 activation. Conversely, our structure appears more consistent with a mechanism, in which phosphorylation would switch off

kinase activity. The crystal structure reveals a number of previously unseen features of Cdc7-Dbf4 and how it engages its substrates. Further work will focus on the importance of such features in a more physiological setting (Chapter 6).

Chapter 6. Results 4 - Functional characterisation of Cdc7 mutants in cellula using a conditional knockout cell line

6.1 Aims

Having gained significant insights into the structural features of Cdc7 required for its activation and substrate recognition, a logical next step in the project was to establish the relevance of these features in a more physiological setting. Until this point, the functionality of Cdc7-Dbf4 mutants could only be tested using *in vitro* kinase assays. Such experiments gave valuable insight into whether the kinase, acting in isolation, is capable of phosphorylating a simple Mcm-derived substrate peptide. However, in the context of a cell, there are numerous further requirements for Cdc7 function, such as the ability to interact with target protein-DNA complexes and numerous further layers of regulation beyond Dbf4 binding.

In this chapter the production of a conditional *CDC7* knockout cell, which can be used as a system to study the kinase mutants *in cellula* is presented. Through utilising retroviral transduction and the bacteriophage P1 Cre-Lox recombination system, Cdc7 mutants could be expressed in the cells in the absence of the WT kinase to test their ability to support Mcm phosphorylation and cell cycle progression. This system is used to validate the importance of structural features described in Chapter 4 and 5, and investigate the possible roles of Cdc7 KI2 and KI3, which appear to be largely dispensable for the *in vitro* activity of the kinase.

6.2 Production of an inducible *CDC7* knockout cell line

To investigate the effect of mutations in *CDC7* on the cell cycle, a conditional knockout cell line, with the endogenous *CDC7* alleles disrupted was created. The

cell line was supported by an artificial *CDC7* allele flanked with LoxP recombination sites. This design allows removal of the artificial allele by the action of bacteriophage P1 Cre recombinase, which can be conveniently expressed from an adenoviral vector (Anton and Graham, 1995). Epithelial HT1080 cells (Rasheed et al., 1974) were chosen as the parental cell line because of the ease of maintenance, high permissivity to adenoviral vectors and the diploid karyotype.

6.2.1 Production of the lentiviral vector containing expressing mCherry-tagged Cdc7 Cdc7

To monitor expression of the artificial (excisable) *CDC7* allele in the presence of endogenous Cdc7, it was modified by an N-terminal fusion to mCherry. This allowed the expression of the transgene to be detected easily and simultaneously with endogenous Cdc7 by Western blotting. It also provided a simple way of monitoring the loss of mCherry-Cdc7 expression upon excision of the artificial *CDC7* allele. The N-terminal residues of Cdc7 were predicted to be disordered, and are not essential for the kinase function *in vitro* (Hughes et al., 2012). Therefore, the addition of an N-terminal tag was deemed less likely to affect the structure and the activity of the kinase. To ensure that the excisable allele is stable in the presence of CRISPR-Cas9 constructs targeting endogenous *CDC7* alleles, the artificial gene was made using a codon-shuffled sequence, originally designed for expression of Cdc7 in bacteria (Hughes et al., unpublished). Full description of the production of the cell line can be found in section 2.2.2.

The mCherry tag adds 29 kDa to the mass of the fusion protein allowing simultaneous detection of the endogenous and ectopically expressed proteins. Approximately 100 clones were tested for mCherry-Cdc7 expression, of which many appeared to overexpress the construct relative to endogenous Cdc7**.** Clones with relatively low expression were expanded and re-blotted (Figure 6-1A), of which one (clone 110), with the closest expression level to endogenous Cdc7 was selected for the next steps in cell line production. Before disruption of the endogenous *CDC7,* the cell line was tested to ensure the LoxP sites had not been

mutated, and that the mCherry-Cdc7 expression could be successfully depleted upon adenoviral transduction (Figure 6-1B). This new cell line was designated HT1080 mCherry-Cdc7(LoxP) and used for further steps in production of the conditional knockout system.

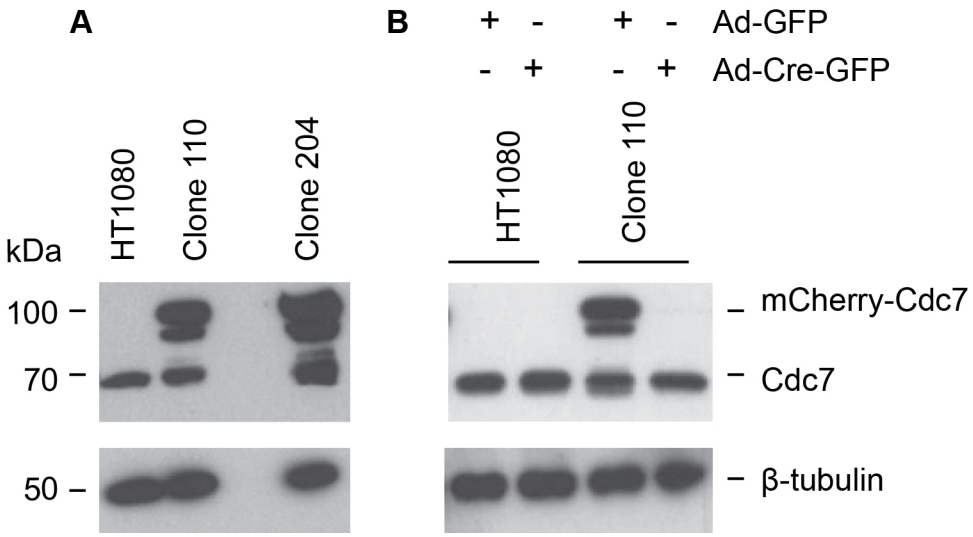


Figure 6-1 Expression and depletion of mCherry-Cdc7 in HT1080 cells

A) Cells were infected with a lentiviral vector expressing pWPT-mCherry-Cdc7 and diluted for single colonies before expanding for several weeks. Individual clones were harvested and lysed prior to Western blotting.
B) Cells from clone 110 were infected with Ad-GFP or Ad-Cre-GFP and grown for 3 d before harvesting and lysis prior to Western blotting. All blots contained 10 µg of cell lysate per well and were probed with an anti-Cdc7 antibody and an anti-β-tubulin antibody as a loading control.

6.2.2 Disruption of the endogenous *CDC7* alleles in HT1080 mCherry-Cdc7(LoxP) cells

To disrupt the endogenous *CDC7* gene the CRISPR/Cas9 system was used. CRISPR is a genome editing technique that has been developed from an adaptive bacterial immune system (Barrangou et al., 2007). The CRISPR system requires two components in the form of a guide RNA and a CRISPR associated endonuclease in the form of Cas9. The guide RNA (gRNA) is a sequence comprising a 'scaffold sequence' required for Cas9 binding and a 20-nucleotide sequence complimentary to the genome to be altered (Jinek et al., 2012). In order to produce a successful knockout, the gRNA must be upstream of a protospacer adjacent motif (PAM) required for Cas9 binding and contain a unique guide sequence. The PAM sequence is essential to binding and specific to each endonuclease (Sternberg et al., 2014). The gRNA sequence can be cloned into specially designed vectors that are then transfected into target cells. Upon expression the Cas9 forms a riboprotein complex with the gRNA, inducing a conformational change in Cas9 that allows it to bind DNA. The complex is then able to bind any PAM sequence in the genome. The affinity for sites in the genome is dictated by the guide sequence and influences how likely Cas9 is to cut at that position in the genome. Cas9 will only cut when the guide sequence is successfully annealed to the genomic sequence. Successful binding leads to a further conformational change in Cas9 that allows it to cut both sides of the DNA resulting in a double strand break 3-4 bp upstream of the PAM sequence. A knockout is achieved when the break is repaired by the error-prone non-homologous end joining mechanism (Jinek et al., 2012). Transfection of the Cas9-gRNA constructs leads to a mixed population of cells containing a range of indels or more complex chromosomal rearrangements. Individual clones can be isolated by limited dilution of the heterogeneous population for detailed genetic and phenotypic characterisations. After transfecting with CRISPR/Cas9 constructs cells were diluted for single colonies, which, after expansion, were tested for Cdc7 expression by Western blotting. Western blots indicated a lack of endogenous Cdc7 expression in the clone 47 (Figure 6-2A and B). The chromosomal region targeted by the gRNA construct in this clone was PCR-amplified, and the product was

cloned into the Topo-TA cloning vector (Life Technologies). Sequencing of 4 individual clones revealed a single frame shift mutation (Figure 6-2C). It seems unlikely that both alleles in the cell clone sustained identical mutations. Given that no WT sequence was detected in this clone, the second allele may have suffered a sufficiently large deletion to prevent amplification with the primers used for PCR amplification of the genomic region.This clone was designated HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) and was used for the complementation experiments.

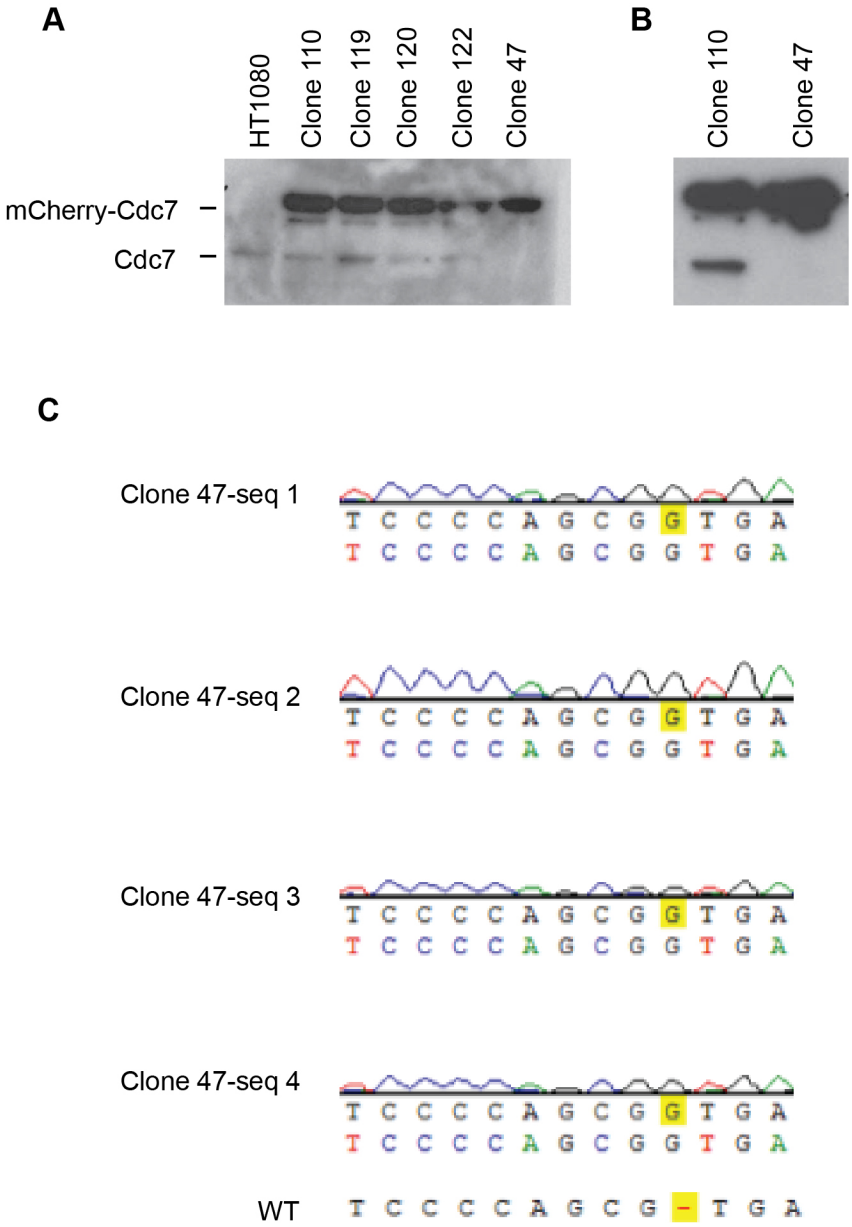


Figure 6-2 Identification of a double Cdc7 knockout after transfection of HT1080 mCherry-Cdc7(LoxP) cells with CRISPR/Cas9 constructs

A) HT1080-mCherry-Cdc7(LoxP) cells were transfected with CRISPR/Cas9 constructs designed to disrupt the *CDC7* gene. Cells were diluted for single colonies before expanding for several weeks. Individual clones were harvested and lysed prior to Western blotting using an anti-Cdc7 antibody.
B) Clones 110 and 47 were blotted for a second time to confirm the full knockout of endogenous Cdc7 expression in clone 47.
C) The region of the genome targeted for disruption by the gRNA was PCR amplified and cloned into a Topo-TA cloning vector. Sequencing of 4 clones using the BigDye Terminator 3.1 sequencing kit revealed that a G-insertion (highlighted in yellow) lead to a frame shift in the endogenous *CDC7* gene. Sequencing was carried out at the in-house sequencing facility.

6.2.3 Expression of FLAG-tagged Cdc7 mutants in *CDC7(-/-)* mCherry-Cdc7 cells

DNA constructs for expression of WT and mutant versions of FLAG-Cdc7 were made on the basis of the pBabe-puro vector (Morgenstern and Land, 1990) (2.2.2). The vector contains a puromycin resistance gene allowing for quick selection of stably transduced cells. Retroviral vector particles were produced by co-transfection of pBabe-FLAG-Cdc7 vectors with the packaging vectors pCG-GagPol and pMDG (Naldini et al., 1996) into HEK-293T cells. Cell-free supernatants containing vector particles were used to infect the *CDC7(-/-)* mCherry-Cdc7(LoxP) cells, and stably transduced cells were selected in the presence of puromycin. Western blot analysis was used to verify expression of FLAG-Cdc7 (Figure 6-3). Near-endogenous expression was achieved for all but the constructs carrying the Δ2aq deletion in KI2, which were all highly over expressed. Attempts were made to reduce the expression of these constructs by detuning the Kozak consensus sequence and using alternative expression vectors. However, Δ2aq Cdc7 was overexpressed in all cases (data not shown), and the original overexpressing cell line was used for further experiments.

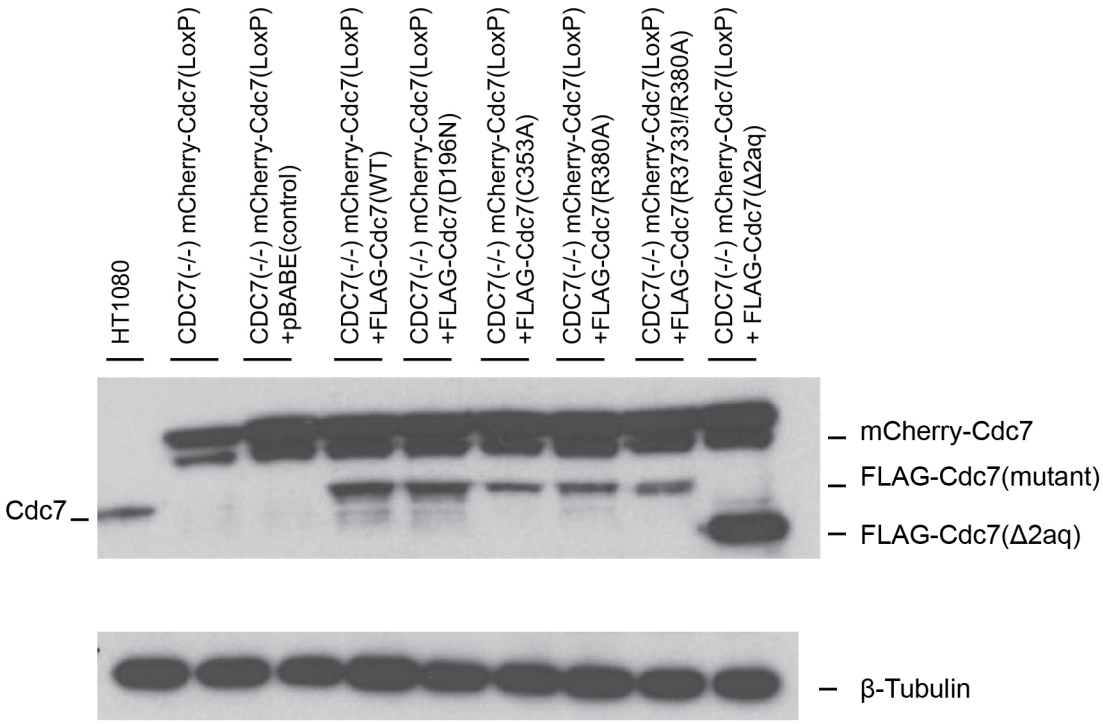


Figure 6-3 Expression of FLAG-Cdc7 mutants in CDC7(-/-) mCherry-Cdc7(LoxP) cells after infection with pBabe(puro) expression constructs

CDC7(-/-) mCherry-Cdc7 (LoxP) cells were stably transduced with retroviral vectors expressing a range of FLAG-Cdc7 variants. Total cell extracts (10 μg protein per lane) were analysed by Western blotting with an anti-Cdc7 (top) and an anti-β-tubulin (bottom) antibody.

6.2.4 Depletion of mCherry-Cdc7 following the treatment of CDC7(-/-) mCherry-Cdc7 cells with Ad-GFP-Cre

To excise the ectopic mCherry-Cdc7(LoxP) allele, the cells were incubated with adenoviral vector particles expressing P1 phage Cre recombinase and GFP from two separate promoters (Ad-GFP-Cre, Vector Biolabs). Cre recombinase catalyses site-specific recombination between two 34-bp DNA sequences called LoxP sites. Recombination between a pair of directly repeated LoxP sites leads to excision of the intervening sequence, leaving a single LoxP sites (Nagy, 2000). The adenoviral system used in this assay leads to excision of the transgene without replication of the virus and subsequent cell death. This is due to a deletion in the *E1* gene in the

vector genome, which prevents virus replication in normal cells. Adenoviruses are able to infect a wide range of cells making them an excellent vector for protein expression in human cells (Anton and Graham, 1995). As a control, adenoviral vector expressing GFP alone was used (Ad-GFP). The amount of virus is expressed in plaque forming units (PFUs), determined by infection of cells permissive for lytic replication of the vector.

Cells were harvested and 70,000 cells infected with $2.5x10^6$ or $5x10^6$ PFU/ml Ad-GFP-Cre or Ad-GFP in 12-well plates for 24 h. At this stage, all cells expressed GFP, as judged by epifluorescence microscopy (data not shown). Cells were expanded into 10-cm dishes for 3 days prior to lysis and analysis by Western blotting with antibodies to Cdc7 and β-tubulin, as a loading control (Figure 6-4).

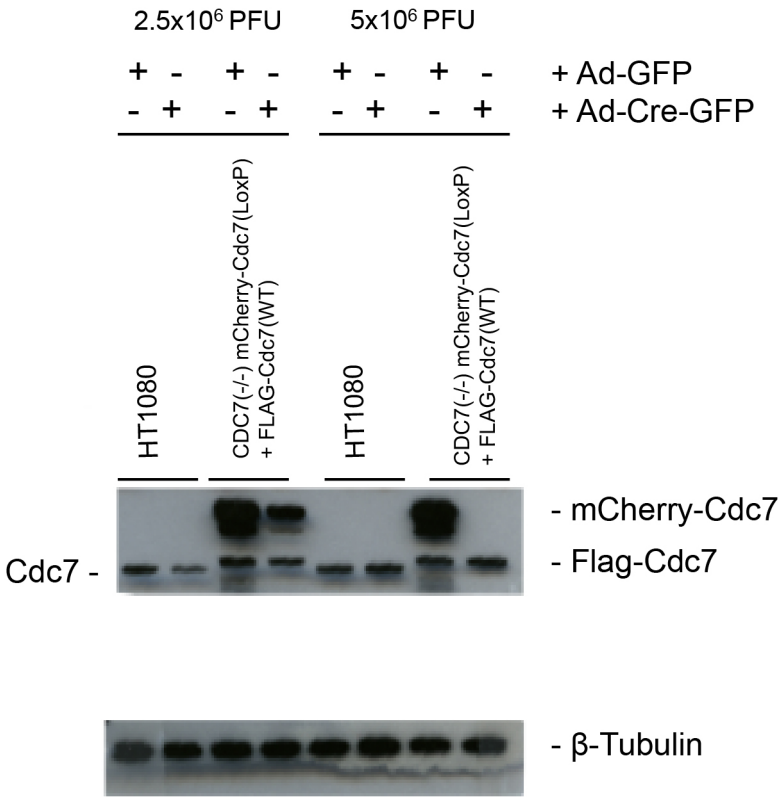


Figure 6-4 Removal of mCherry-Cdc7(LoxP) by Cre recombinase

HT1080 and *CDC7(-/-)* mCherry-Cdc7 (LoxP) cells were infected with $2.5x10^6$ and $5x10^6$ PFUs of Ad-GFP or Ad-Cre-GFP and grown for 3 d prior to harvesting and lysis; 10 µg of total protein was separated in SDS-PAGE gels and analysed by Western blotting using anti-Cdc7 and anti-β-tubulin antibodies.

6.3 Complementation of cell cycle profiles and Mcm2 phosphorylation by mutants of Cdc7

For the cell cycle assays, each cell line was infected with $5x10^6$ PFU of Ad-GFP or Ad-GFP-Cre and expanded for 5 d prior to cell cycle analysis and Western blotting with antibodies specific to Mcm2 phosphorylated at Ser40. Excision of the transgene was also verified using an anti-Cdcy antibody (Figure 6-5).

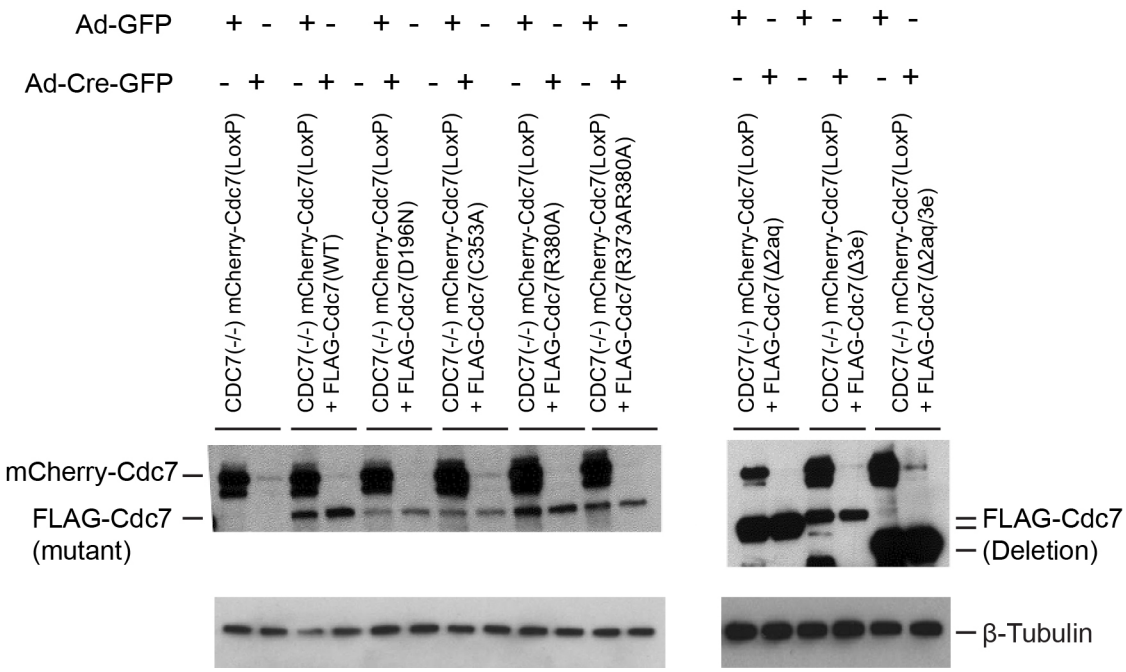


Figure 6-5 Depletion of mCherry-Cdc7 following infection with Ad-GFP-Cre

Cells infected with Ad-GFP or Ad-GFP-Cre were expanded for 5 days prior to Western blotting with anti-Cdc7 antibodies; each lane contained 10 µg of total cell extract.

6.3.1 Cell cycle assay following depletion of mCherry-Cdc7 in the conditional knockout cell line

Because Cdc7 is essential for cell proliferation, its depletion or inhibition leads to a cell cycle arrest and eventual cell death (Ito et al., 2012). Furthermore, infection of cells with adenoviral vectors or expression of Cre recombinase could in principle affect cell growth. For this reason, the conditional knockout *CDC7(-/-)* mCherry-Cdc7(LoxP) cells transduced with pBabe-FLAG-Cdc7(WT) or pBabe-FLAG-Cdc7(D196N) were used to validate downstream analyses. Cdc7 Asp196 is part of the essential DFG kinase motif, mutations of which are incompatible with the kinase activity. Cells were expanded for 5 d prior to harvesting and cell cycle analyses by flow cytometry (2.5). At this time, cell lysates were also prepared for Western blotting with phosphospecific antibodies to detect Cdc7-specific phosphorylation of Mcm2 Ser40.

Addition of the FLAG-Cdc7(WT) construct displayed cell cycle profiles similar to the GFP expressing cells upon mCherry-Cdc7 removal (Figure 6-6A). The Cre recombinase expressing cells also maintained normal levels of Mcm2 Ser40 phosphorylation, when compared to the GFP transduced control cells (Figure 6-7). Visualising the cells 7 days after excision also revealed that cells proliferate at a similar rate to control cells after transgene excision, and these cells exhibit a normal morphology (Figure 6-6A). Therefore, as expected, FLAG-Cdc7(WT) is able to support normal cell cycle. In contrast, upon infection with Ad-GFP-Cre, the cells expressing FLAG-Cdc7(D196N) displayed an accumulation in G2/M with concomitant reductions in S and G1 populations (Figure 6-6B) and a near total loss of Mcm2 Ser40 phosphorylation (Figure 6-7). Considerably fewer cells were seen when visualised after 7 days and many of the remaining cells exhibited an unusual morphology consistent with a failure to perform successful mitosis. As expected, the cells maintained normal cell cycle profile upon infection with the control Ad-GFP vector (Figure 6-6B).

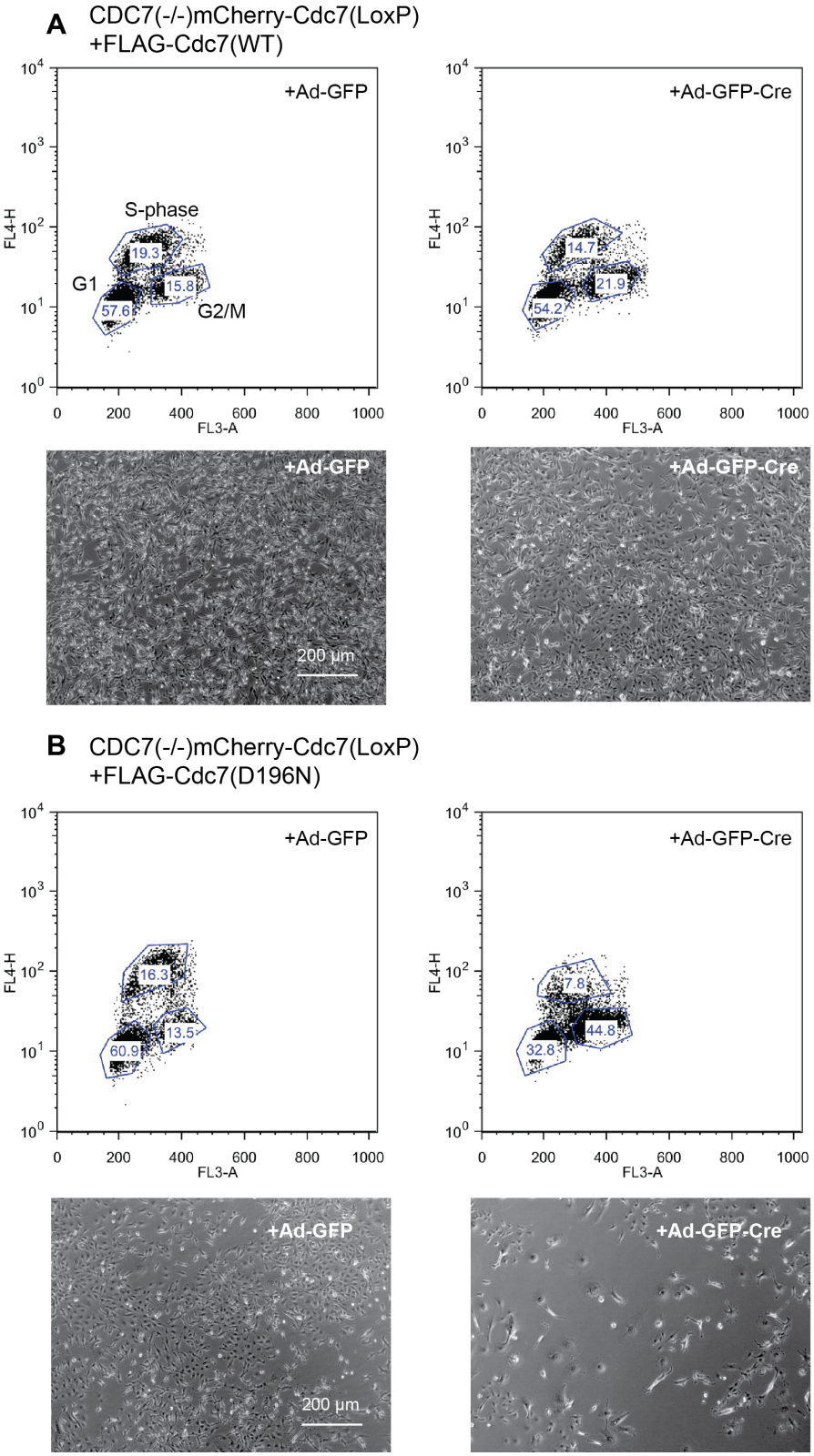


Figure 6-6 Cell cycle analysis of *CDC7(-/-)* mCherry-Cdc7(LoxP) cells expressing WT or kinase dead D196N FLAG-Cdc7 after mCherry-Cdc7 excision

A) Cell cycle profile and photographs of cells transduced containing FLAG-Cdc7(WT) B) Cell cycle profile and photographs of cells transduced containing

FLAG-Cdc7(D196N). Cells were stained for total DNA with propidium idodide (detected in FL3 channel) and nascent actively replicating DNA (Alexa Fluor 647, FL4 channel). Cells were then analysed by flow cytometry (for details see Materials and methods 2.5) Cells were analysed 5 days post infection with Ad-GFP or Ad-Cre-GFP. Cells were photographed 7 days post infection.

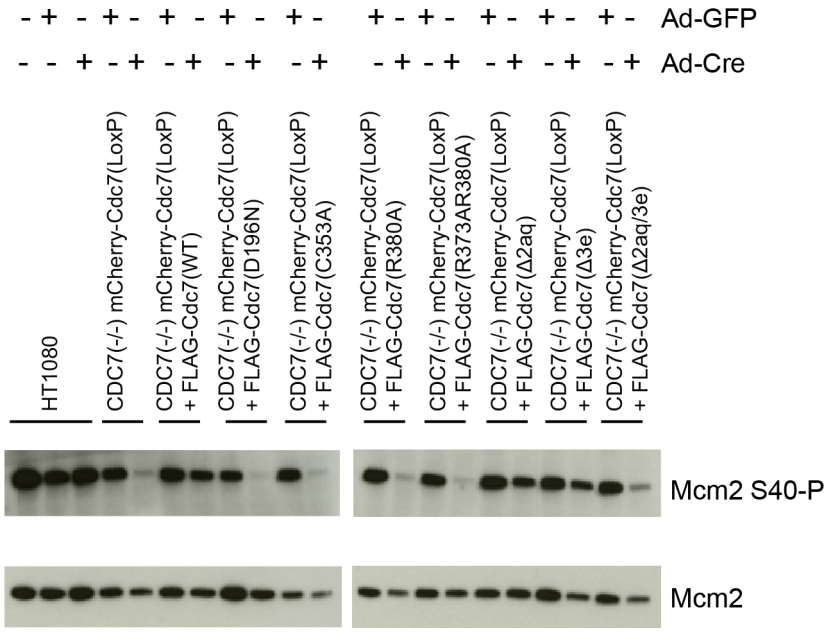


Figure 6-7 Western blot of Mcm2-S40 phosphorylation

CDC7(-/-) mCherry-Cdc7(LoxP) cells expressing FLAG-Cdc7 and its mutant forms (as labelled) were infected with Ad-GFP or Ad-GFP-Cre and allowed to expanded for 5 d before harvesting and Western blotting with anti-Mcm2 Ser40 and anti-Mcm2 antibodies; each lane contained 10 µg of total cell protein.

6.3.2 Point mutations of the Cdc7 Zn-binding domain, Arg373 or Arg380 affect Mcm2 phosphorylation and cell cycle progression

The experiments described above established that mCherry-Cdc7 fusion can be successfully depleted from *CDC7(-/-)* mCherry-Cdc7(LoxP) cells, and that the survival of the cells upon the depletion depends on complementation with a functional Cdc7 variant. The system was then applied to test functionality of a range of FLAG-Cdc7 mutants (Figure 6-5). Cell cycle profiles and the

phosphorylation status of Mcm2 Ser40 were monitored following infection of the cell lines with Ad-GFP-Cre, relative to control cells infected with Ad-GFP.

Mutation of the invariant Cdc7 Cys353 involved in metal coordination within the Zn-binding domain (C353A) lead to an accumulation of cells in G2/M phase of the cell cycle (Figure 6-8), albeit less pronounced than with the D196N mutant. Near total loss of Mcm2 Ser40 phosphorylation was also observed (Figure 6-7). The difference in cell number and morphology between Cre and control treated cells is also similar to that of the D196N mutant, with a significant reduction in surviving cells and a significant number of cells exhibiting unusual morphology. This is in excellent agreement with *in vitro* kinase assays (Figure 3-5) and confirms the critical importance of the Zn-binding domain in a physiological setting.

Mutation of Arg380 to Ala leads to an accumulation of cells in G2/M and a significant decrease in Mcm2 Ser40 phosphorylation (Figure 6-9 and 6-7). In the case of the double mutant R373A/R380A, a more pronounced accumulation of cells in G2/M of the cell cycle was observed, alongside the same significant decrease in steady-state levels of MCM2 Ser40 phosphorylation (Figure 6-9 and 6-7)). The number of surviving cells and cell morphology of Cre treated cells also mirror that of the D196N mutant cells line. However, in the case of the single R380A mutant this is much less pronounced than the double mutant. This data is in excellent agreement with *in vitro* kinase data, in which the double mutant had significantly reduced activity relative to R380A alone (Figure 3-9). The ability of the R380A single mutant to survive to a greater extent than other mutants may be due to its ability to still successfully target a number of other Mcm2-7 phosphorylation sites as well as its slightly increased ability to phosphorylate Mcm2-S40. To test this, it would be possible to use antibodies specific to phosphorylation of other Mcm2-7 target sites to see if the R380A mutant significantly affects them.

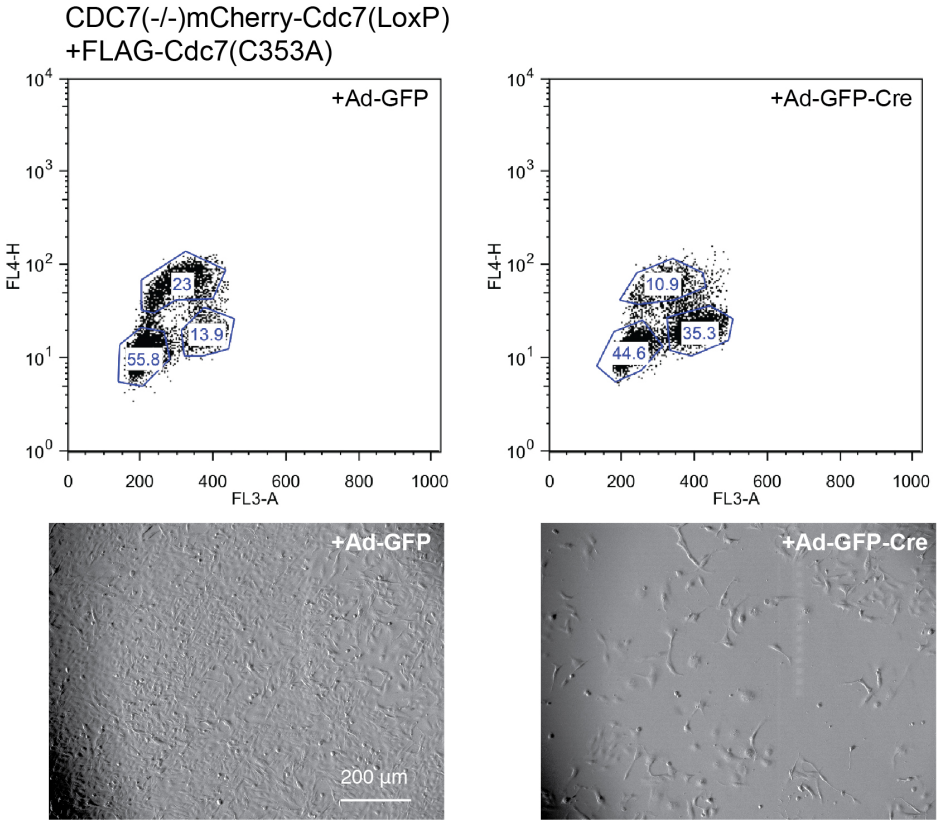


Figure 6-8 Cell cycle analysis of *CDC7(-/-)* mCherry-Cdc7(LoxP) cells expressing FLAG-Cdc7 (C353A) infected with Ad-GFP-Cre or Ad-GFP

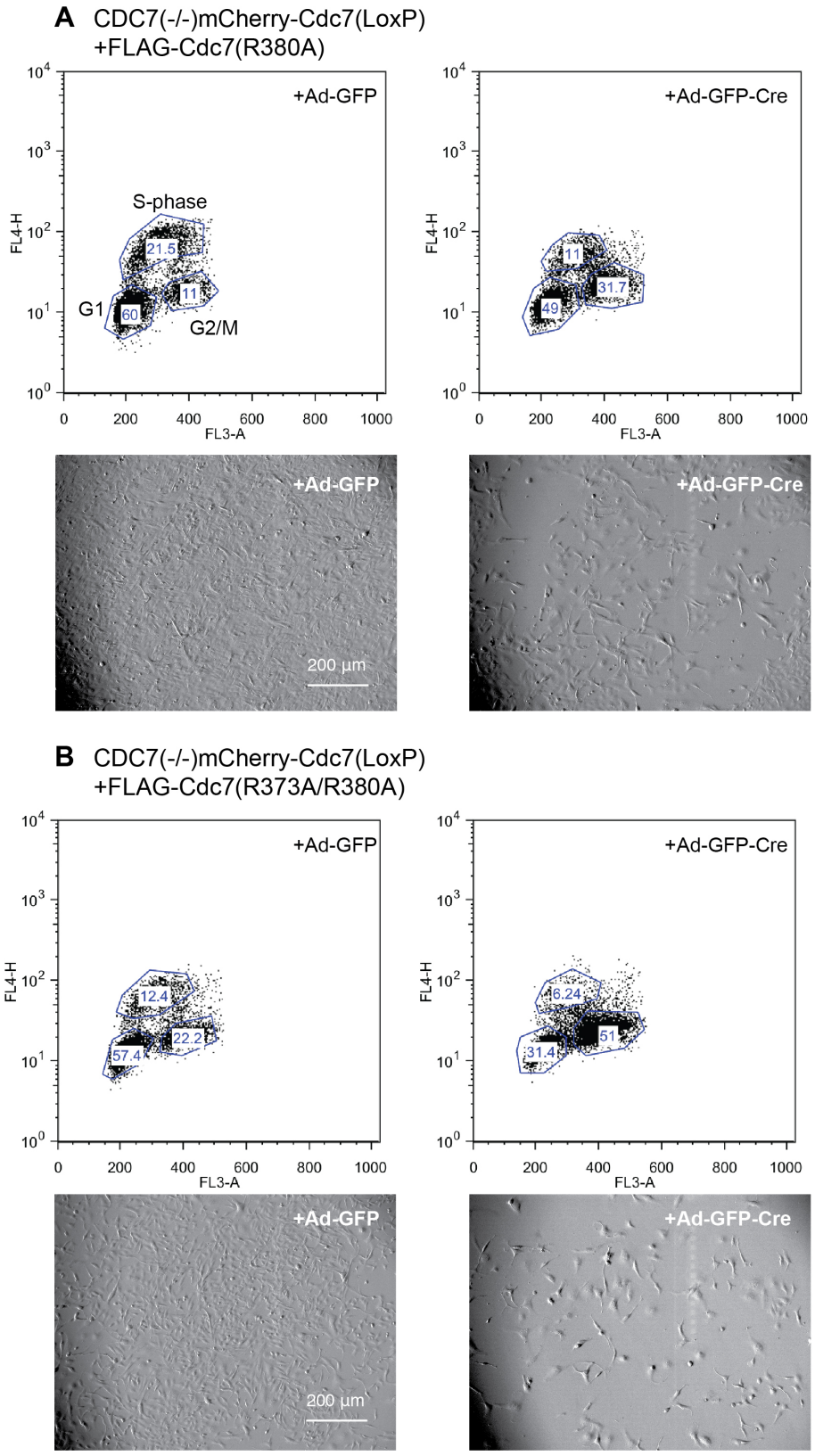


Figure 6-9 Cell cycle analysis of *CDC7(-/-)* mCherry-Cdc7(LoxP) cells expressing R380A or R373A/R380A mutants of FLAG-Cdc7 after infection with Ad-GFP-Cre or Ad-GFP

6.3.3 Deletions in KI2 and KI3 have a minimal affect on Mcm2 phosphorylation and cell cycle progression

To determine if the regions of Cdc7 KI2 and KI3 deleted in the crystallized construct are essential for its functionality *in cellula*, we tested the ability of Δ2aq and Δ3e mutants of Flag-Cdc7 to complement the depletion of mCherry-Cdc7 in the conditional knockout cell line. Surprisingly, despite its extent, the Δ2aq deletion had relatively modest effect, with 13% of the cells undergoing active S-phase (Figure 6-10A). Consistent with these observations, the steady-state levels of Mcm2 Ser40 phosphorylation were comparable to the control condition (Figure 6-7); moreover, the cells continued to proliferate for at least 2 passages. Differences in cell number and morphology between the Cre treated and control cells are also minimal when compared to the previous mutants. Although these results indicate that a large section of KI2 may be dispensable both *in vitro* and *in cellula*, overexpression of the Δ2aq mutant may compensate for the defect in its activity.

FLAG-Cdc7(Δ3e), harboring the KI3 deletion used for crystallization was capable of supporting normal cell cycle, with 18.5% of the cells in active S-phase (Figure 6-10B), and normal steady-state levels of Mcm2 Ser40 phosphorylation (Figure 6-7). In contrast to Δ2aq, this mutant was expressed at near endogenous levels, arguing that the removed segment of KI3 may be indeed dispensable for cell cycle progression. The number and morphology of cells observed at 7 days, in which the control and Cre treated cells appear similar. This is in agreement with *in vitro* kinase assays (Figure 3-6).

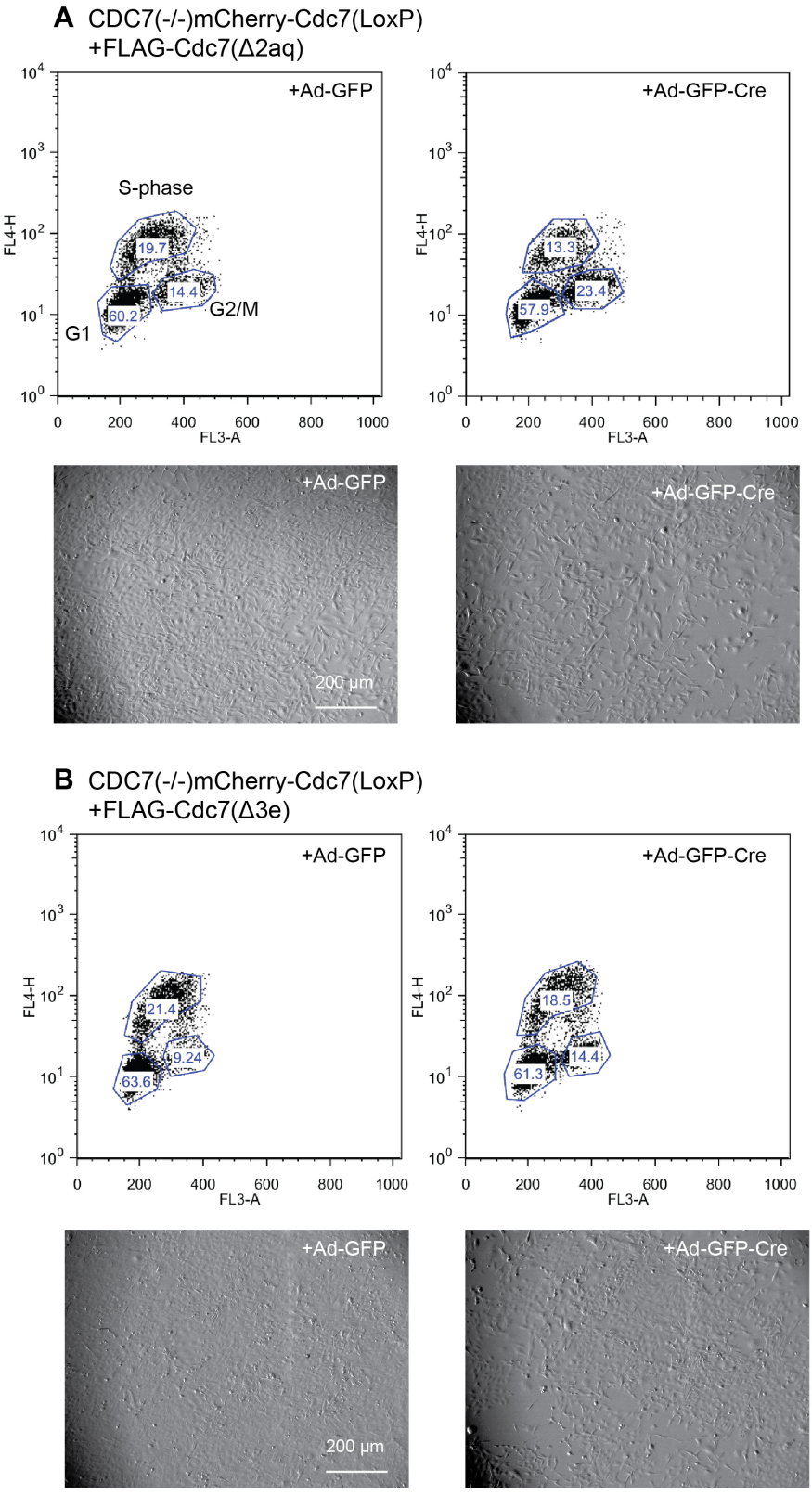


Figure 6-10 Cell cycle analysis of *CDC7(-/-)* mCherry-Cdc7(LoxP) cells expressing Δ2aq orΔ3e deletion mutants of FLAG-Cdc7 after mCherry-Cdc7 excision

6.4 Conclusions

In this chapter the production of a conditional Cdc7 knockout cell line is described for investigating the importance of a number of structural features of Cdc7 in a physiological context. By introduction of a tagged Cdc7 transgene flanked by LoxP sites and subsequent removal of endogenous Cdc7 expression, it was possible to construct an assay system, in which mutants of Cdc7 could be introduced into the cell and tested for their ability to support cell cycle progression. Point mutations in structural features described in Chapter 4 and Chapter 5 produced outcomes in agreement with previously used *in vitro* kinase assays. Disruption of the Zn-binding domain through mutation of the invariant Cys353 residue almost phenocopied the kinase dead mutant D196N, confirming importance of the Zn-binding domain for the core Cdc7 functions in the cell. Mutation of the conserved residues Arg373 and Arg380 also lead to reduced Mcm2 Ser40 phosphorylation and a reduction in cell cycle progression, with a more pronounced phenotype when both residues were mutated simultaneously. These results further highlight the importance of the interactions between Cdc7 and the P+1 residue of the target site in ensuring full kinase activity.

The Δ2aq and Δ3e FLAG-Cdc7 mutants were capable of supporting cell proliferation with minimal effects on steady state levels of Mcm2 Ser40 phosphorylation. Therefore, KI2 and KI3 residues deleted in the crystallized construct may not be strictly essential for the cell cycle related functions of Cdc7. However, it is important to note that overexpression of FLAG-Cdc7(Δ2aq) mutant could compensate a partial loss of function.

As deletions in KI2 appeared to have limited affect on cell cycle progression, it was hypothesized that this region may be important for other roles Cdc7 plays in the cell. As discussed in the introduction section, Cdc7 interacts with a number of different proteins and acts in various capacities in the cell, including the DNA damage response and meiosis. (Tsuji et al., 2008, Sheu and Stillman, 2010, Lee et al., 2012). This represents a future and potentially interesting avenue for investigation. However it is important to consider such conclusions as preliminary as the

presented data is representative of a single experiment. To fully confirm the effect of such mutants it is necessary to perform the experiment a number of times to account for the inherent variability in assays utilizing live cells. While there appears to be a reduction in active S-phase for a number of mutants, these reductions are often relatively small so replicates are required to confirm their significance. However, the accumulation of cells in G2/M, when combined with the difference in the number of cells and cellular morphology of a number of Cre treated cells does suggest an effect on cell viability and cell cycle progression for a number of mutants identified in the obtained structures. Interestingly, the accumulation of cells in G2/M is not in agreement with previous studies, which suggest cells with a functional allele of p53 should arrest and die during an abberant S-phase (Ito et al., 2012). It is possible that accumulation may be an artifact of the system due to the very slow degradation of the mCherry-Cdc7 transgene. Development of a system with a faster knockout or production of a similar system in cell lines both positive and negative for p53 could shed light onto this unexpected outcome. A further important avenue of investigation would be to ascertain whether the cells have simply ceased to proliferate or have died through apoptosis. This could be investigated with simple apoptosis assays in which caspase activity can be measured.

Collectively, the results described in this chapter suggest that the features identified in the crystal structure of Cdc7 may be physiologically important. The conditional Cdc7 knockout cell line represents a potential tool for understanding the various currently unexplained features of Cdc7 in the future. In particular, the potential roles of the extended KI2 in other cellular processes, such as the response to replication stress, warrant further investigations.

Chapter 7. Discussion

7.1 Introduction

The S-phase kinase Cdc7-Dbf4 has a range of critical cellular functions and has attracted a lot of attention in recent years, not least because of its potential as a target for cancer therapeutics (Sawa and Masai, 2009). Overexpression of Cdc7 and Dbf4 is common in transformed cell lines and primary cancer samples (Nambiar et al., 2007, Bonte et al., 2008, Kaufmann et al., 2008, Clarke et al., 2009, Kulkarni et al., 2009, Hou et al., 2012, Cheng et al., 2013, Melling et al., 2015, Ghatalia et al., 2016). Moreover, selective inhibition was claimed to kill transformed cells, while leaving regular tissues unharmed (Ito et al., 2012). A number of Cdc7 inhibitors have already been developed with some success (Zhao et al., 2009, Koltun et al., 2012, Sasi et al., 2014, Li et al., 2015b), however, the design of future more potent and selective inhibitors could be greatly helped by better understanding of its structure.

Cdc7-Dbf4 structural data has so far been mostly limited to conserved domains common to all kinases and a small fragment of Dbf4 minimally required for kinase activation (Hughes et al., 2012). The original structure revealed that Dbf4 uses its conserved motif-M to tether itself to the C-lobe of kinase subunit and stabilized the Cdc7 αC in its active “in” conformation using motif-C. However, deletions in the kinase insert sequences of Cdc7 (KI2 and KI3) were required for crystallisation. Deletions in KI2 also lead to a significant decrease in *in vitro* kinase activity, suggesting essential features of the kinase were missing from the construct (Hughes et al., 2012).

Until now a structure of Cdc7-Dbf4 construct with activity similar to that of WT heterodimer has not been reported, and no structural information on how the kinase engages its target substrates is available. The main focus of this project was to fill this gap in our understanding of Cdc7-Dbf4 by determining crystal structures of the kinase with a more complete KI2 and hopefully improved levels of activity.

Such structures are presented in chapters 4 and 5, with the rationale for the construct design described in chapter 3. Further to this, the structure was validated both *in vitro* and *in cellula* using a conditional knockout cell line designed as part of the study (chapter 6). This not only confirmed the importance of the structural features of Cdc7-Dbf4 revealed in this study, but also opened an avenue for functional analyses of the poorly understood insert sequences of Cdc7.

7.2 Key findings

7.2.1 New insights into the structure of Cdc7-Dbf4

A key finding of this study was the identification of a novel Zn-binding domain in Cdc7 KI2 (Figure 4-5). A single Zn atom is coordinated by four Cys residues, which are invariant across all metazoan orthologues of Cdc7. This feature was shown to be critical for the kinase activity *in vitro* and essential for the ability of the kinase to support cell proliferation. Mechanistically, the interaction between the Zn-binding domain and Dbf4 pins back the activation loop of Cdc7, which is extended by KI2, fully opening up the active site of the kinase (Figure 3-5, 4-4 and 6-8). The pinning back of the loop creates a substrate peptide-binding platform that was partially occluded by a disordered portion of KI2 in the previously reported crystal structure. Interestingly, the invariant KI2 Cys residues that coordinate Zn are not conserved in lower eukaryotes, suggesting an alternative mechanism for maintaining an open conformation in yeast systems. Structures of Cdc7-Dbf4 were obtained bound to a potent Cdc7 inhibitor (XL413) and an ATP mimic (ADP-BeF_3^-). In the latter structure, the BeF_3 moiety that mimics the γ-phosphate of the nucleotide could be visualised alongside the two Mg^{2+} ions involved in the catalysis.

7.2.2 Substrate binding of Cdc7-Dbf4

The fully open conformation of Cdc7-Dbf4 in the new structure allowed for crystallisation of the heterodimer in complex with an Mcm2-derived substrate peptide that had been covalently attached to ATPγS (referred to as bi-substrate ligand). This approach allowed successful co-crystallization of the kinase with an analogue of a substrate peptide, which could be visualised in the structure, despite the apparent hydrolysis of the chimeric bond. The newly obtained structure

confirmed the importance of the conserved Arg373 and Arg380 residues that had been hypothesised to interact with the negatively charged P+1 residue of the substrate (Figure 5-7). The importance of this feature was again validated both *in vitro* and *in cellula* highlighting their importance for kinase function (Figure 4-9, Figure 6-9). Further to this, an unexpected role for the P+4 residue of the peptide was observed in the structure. This residue, corresponding to Mcm2 Arg44, is involved in a network of interactions with both Cdc7 and Dbf4 in a binding cleft formed between the kinase and activating subunit. Intriguingly, the P+4 position of such a pre-phosphorylated Cdc7 target site also represents the P+3 residue of the consensus sequence of CDK2 that primes this site for Cdc7 targeting (Stevenson-Lindert et al., 2003). This observation suggests that Cdc7 is optimized to act in tandem with CDKs regulating DNA replication. However, while this pocket makes favourable interactions with the given substrate, it is important to recognise that the P+4 position of a Cdc7 target site can be occupied by a number of different residues, and the pocket therefore must therefore be able to accommodate a range of amino acids side chains. This was confirmed *in vitro* with peptides containing substitutions at the P+4 position (Figure 5-9).

7.2.3 A possible role for Thr376 in kinase regulation

Many protein kinases require phosphorylation of their activation loops at one or more Thr, Ser, and/or Tyr residue(s) for full activation (Adams, 2003). These phosphorylation events induce considerable conformational rearrangements, such as those seen in CDK2-CyclinA, where phosphorylation of CDK2 Thr160 induces the formation of a hydrogen-bonding network resulting in opening of the catalytic cleft (Russo et al., 1996, Brown et al., 1999a). In a previous study, it was proposed that Cdc7 may be regulated by a similar mechanism, whereby conserved Cdc7 Thr376 residue was proposed to act as an activation Thr (Masai et al., 2000). However, structural data obtained in this study argue that the location of Thr376 would be incompatible with an active kinase conformation due to the close proximity to the negative charges of the γ-phosphate and the negatively charged P+1 substrate residue. These negative charges would likely lead to charge repulsion that would prevent the binding of the substrate. Furthermore, the

interaction between KI2 Zn-binding domain and Dbf4 that orders the activation loop of Cdc7 appears to be functionally analogous to the effect of activation loop phosphorylation in other kinases.

Intriguingly, the location of Cdc7 Thr376 suggests that its phosphorylation could act as an off switch for Cdc7, and removal of the phosphate by a protein phosphatase could lead to re-activation of the kinase. This idea has received some support in the form of a recent study, in which phosphorylation of numerous sites of Cdc7 (including Thr376) by CDK1 led to its dissociation from chromatin to prevent DNA re-replication in mitosis. Substitution of Cdc7 Thr376 for Ala conferred a reduction in kinase activity *in vitro* (Knockleby et al., 2016). Furthermore, PP1 was implicated as the phosphatase required for removal of the phosphorylation to facilitate re-loading of Cdc7-Dbf4 onto chromatin in the next cell cycle. These data support a model, in which de-phosphorylation of the activation loop of Cdc7 is necessary for the full activity of the kinase. Further to this, our immunoprecipitation experiments and mass spectrometry identified an additional protein phosphatase, PP6 as a binding partner of Cdc7-Dbf4 in cells (data not shown). Intriguingly, this heterotrimeric phosphatase has been implicated in regulation of other kinases as well as roles in the cell cycle and DNA damage response (Stefansson and Brautigan, 2007, Zeng et al., 2010, Zhong et al., 2011, Hosing et al., 2012).

7.2.4 Functional importance of Cdc7 KI2 and KI3

Previous to this study, the only method for assessing Cdc7 mutants in the lab had been through the use of *in vitro* kinase assays. While these types of experiments are informative, they only reveal how well the kinase, acting in isolation, can phosphorylate a simple substrate. In the cell, Cdc7-Dbf4 relies on a number of protein-protein interactions for correct targeting and activation, which may be impossible to recapitulate in an *in vitro* system. Development of an inducible Cdc7 knockout cell line as part of this study allowed for testing of such mutants *in cellula* (Chapter 6), revealing how modifications in Cdc7 may affect a number of its cellular functions. The system was used to further validate point mutations of Cdc7 based on structural features observed in this study, and these experiments were largely in

agreement with what had previously been seen *in vitro.* An important observation from experiments using the conditional knockout cell line was that the extended portions of KI2 and KI3 do not appear to be strictly required for the function of Cdc7 in cell proliferation and maintaining the steady state levels of Mcm2 phosphorylation (Figure 6-7, 6-10). This was particularly surprising in the case of KI2 due to the extent of the deletion and the fact that KI2 extends the activation loop of Cdc7. It is possible that this large, apparently disordered domain mediates interactions between Cdc7 and its functional partners in the cell. For example, KI2 contains the nuclear localisation sequence for Cdc7 and an interaction between this region of the protein and importin-β has been proposed to facilitate Cdc7 entry into the nucleus. Further to this, a nuclear retention sequence is also present in KI2, which is required to maintain the presence of Cdc7 in the nucleus (Kim and Lee, 2006, Kim et al., 2007). It is important to consider that the KI2 deletion mutant of Cdc7 was overexpressed, and this overexpression could have masked a stronger phenotype. Further to this, this data should be considered as preliminary and requires a number of replicates as well as fine-tuning of the system. However, should the lack of requirement for KI2 and KI3 be confirmed, it is possible that the insert sequences are involved in cellular roles of Cdc7 that have yet to be identified. An interesting idea is that the inserts may be involved in the functions of Cdc7-Dbf4 in meiosis and/or DNA damage response, and these represent interesting avenues for future studies.

7.3 Future directions

7.3.1 Structural characterisation of Cdc7-Dbf4

While my work revealed a number of new Cdc7-Dbf4 features that are critical to kinase activation and substrate specificity, substantial deletions in KI2 and KI3 required for crystallisation mean large sections of the kinase along with the bulk of the Dbf4 structure remain uncharacterized. Future work in this respect should focus on attempting to gain structural data on more complete constructs of Cdc7. Despite the lack of predicted secondary structure and sequence conservation in these sections of the protein, their existence across all eukaryotes suggests an important cellular role. It is likely that a structure amenable to crystallisation is only likely to be

achieved in the presence of a binding partner. My attempts to identify such a partner using immunoprecipitation experiments and mass spectrometry highlighted one interesting candidate, PP6 (data not shown). The newly developed cell line in this study represents a new avenue for identifying proteins in the cell that may interact with Cdc7 in a KI2- and/or KI3- dependent manner. Performing similar immunoprecipitation/mass spectrometry experiments utilising deletion mutants of Cdc7 could identify proteins that require either insert to interact. All constructs of Cdc7 utilised in the conditional knock out cell line were made with an N-terminal FLAG tag to facilitate efficient pull down of the mutants with commercially available antibodies.

Attempts at crystallisation of Cdc7-Drf1 in this study were unsuccessful despite the production of constructs with significant activity *in vitro* (Figure 3-8). Future work should look to obtain a structure of this complex, which could reveal features that explain the requirement for two different activation subunits in the cell. Constructs in this study were designed solely based on homology to Dbf4 and secondary structure predictions. Limited proteolysis of a full Cdc7-Drf1 complex should be performed to better inform a construct design more relevant to the complex. It seems likely that Cdc7-Dbf4 and Cdc7-Drf1 have some degree of redundancy in their functions (Montagnoli et al., 2002), and as potential cancer therapeutics Cdc7 inhibitors may or may not need to target both types of complexes.

Further to this, the structural differences required for binding of different substrate peptides should be further investigated to see how the binding of substrates with Asp or Glu at P+1 position differs from that of the primed phosphorylation sites as well as how different P+4 residues are accommodated in the identified binding pocket. My attempts to crystallise bi-substrate ligands of alternative Cdc7 target peptides have so far been unsuccessful. This is likely due to hydrolysis of the substrate as seen in the structure (Figure 5-6B). Bi-substrate ligands have been designed in other studies, where the peptide or a small protein fragment is covalently linked to a small molecule inhibitor rather than a nucleotide, sometimes leading to further increases in affinity (Schneider et al., 2005, Stebbins et al., 2011). In principle, this approach allows for a wider range of chemistries for constructing

more stable chimeric ligands, which are less likely to undergo hydrolysis during crystallisation.

Lastly, recent developments in cryo-electron microscopy have allowed the replisome and its components to be visualised at a level of detail that was previously impossible (Li et al., 2015a, Abid Ali et al., 2016). A previous issue of crystallising Cdc7-Dbf4 bound to its main substrate (Mcm2-7) was the requirement that the helicase be loaded on DNA for Cdc7-Dbf4 to show maximum affinity for its substrate (Francis et al., 2009). Potentially, cryo-electron microscopy may allow to obtain high-resolution structures of replication origin-loaded Mcm2-7 in complex with the kinase.

7.3.2 Phosphorylation as a mechanism of Cdc7-Dbf4 regulation

Data presented in this thesis suggest a role for phosphorylation of Thr376 of Cdc7 in regulation of kinase activity. Analysis of Thr376 phosphorylation throughout the cell cycle was complicated by the inability to reproducibly detect the phosphorylated peptide by mass spectrometry without overexpressing Cdc7. In this study, a modest reduction in Thr376 phosphorylation was observed in S-phase synchronised cells relative to asynchronous (data not shown). Unfortunately, it was not possible to perform this experiment with endogenous Cdc7 due to the low efficiency of available antibodies for immunoprecipitation of Cdc7 from the cell. However, the conditional knockout cell line produced as part of this study offers the possibility to use efficient affinity tag antibodies to pull down stably expressed FLAG-Cdc7, which may allow for enough material to be obtained for successful analysis. A method has also been optimised by our mass spectrometry facility to facilitate reliable detection of the phosphorylated peptide. These experiments allow for analysis of the phosphorylation state of Thr376 throughout the cell, as well as in the context of various phosphatase knockouts to identify the key player in kinase regulation. The cell line presented in this study could also be used in conjunction with alanine substitution and phosphomimetic mutants of Thr376 to validate the possible roles of this modification in a physiological setting.

The identification of PP6 as a potential interactor is also an intriguing one, and future work should focus on recombinant expression of the complex to validate the interaction *in vitro.* Work in this study suggests that this is not possible in bacteria (data not shown) and expression will likely need to be performed in a eukaryotic system.

7.3.3 Future functional characterisation of Cdc7-Dbf4

The conditional Cdc7 knockout system presented here can be used in a number of ways to validate the importance of KI2 and KI3 in various cellular contexts, and will allow testing of an array of point and deletion mutants of Cdc7, and screen for potential loss of function in cell cycle or recovery from DNA damage response. Future work should therefore focus on detection of phenotypes in cells under genotoxic stress when supported by Cdc7 containing a deletion in an insert sequence. KI2 and KI3 could be important for interactions in the context of the intra S-phase checkpoint and this would be revealed using the presented system when cells have been treated with DNA damaging agents. Phosphoproteomics experiments could also be performed using the conditional knockout cell line complemented with various Cdc7 deletion mutants. Using stable isotope labelled amino acid culture (SILAC), it may be possible to accurately identify and quantify Cdc7 target sites, which are affected by KI2 and 3 deletions. This would not only give further clues as to the functional relevance of the insert sequences, but also could serve as a method for identifying binding partners of Cdc7 for future studies. One limitation of the current system was the overexpression of the mCherry-Cdc7 construct. This meant that depletion of intracellular levels of the fusion protein took several days. This is theorised to be the possible cause of the G2/M accumulation of cells when supported by kinase dead mutants of Cdc7. The expected phenotype for HT1080 cells would be accumulation and cell death in S-phase due to the presence of functioning p53 (Ito et al., 2012). A quick way to confirm whether this phenotype is due to an unexpected change in the HT1080 cells or the slow reduction in mCherry-Cdc7 would be to treat the cells with a potent Cdc7 inhibitor (such as XL413) and see if a rapid inactivation gives a similar phenotype in the newly created cell line.To improve the efficiency of the mCherry-Cdc7 depletion,

the construct could also be tagged with a degron, which can be used to induce rapid degradation of the protein upon treatment with a small molecule such as auxin (Natsume et al., 2016).

7.4 Final conclusions

The data presented in this study provides previously unknown information about the structure and function of Cdc7-Dbf4. The structural bases of its activation and substrate recognition have been substantially clarified. Importance of the features revealed by the crystal structures have been validated *in vitro* and investigated in the context of the cell. However, a number of questions have been raised as part of this study, the biggest of which being the physiological roles of KI2 and KI3. While sequence conservation in theses inserts is limited, and there is little predicted secondary structure, the presence of such insert sequences across eukaryotes suggesting they fulfil important functions in the cell. The inducible Cdc7 knockout cell line produced as part of this study provides a tool to further dissect the roles of these inserts in the cell, as well as offering an avenue for identification of novel interactors and targets of Cdc7-Dbf4. The data presented here provides insights into the structural features required for Cdc7 activation and substrate recognition.

Chapter 8. Appendix

Figure 8-1 Amino acid alignments of distal Cdc7 orthologs (pg172)

The amino acid sequences were from *Homo sapiens* (mammalian), *Gallus gallus* (avian), *Anolis carolinensis* (reptile), *Xenopus laevis* (amphibian), *Danio rerio* (fish), *Danaus plexippus* (insect) and *Saccharomyces cerevisiae* (yeast)**.** Secondary structure elements are indicated atop the alignment. Portions of the protein sequence not present in the crystallized construct are highlighted in gray (1–36, 228–345 and Δ467–533); the kinase insert sequences (black boxes). Invariant kinase residues are highlighted in red and labelled. Features discovered in this study are indicated with red boxes and labelled. Conserved amino acid residues are shown in bold print and those invariant within the alignment are highlighted in yellow (Adapted from Hughes et al. 2012)

Figure 8-2 Amino acid alignments of the motifs-M and C of distal orthologs Dbf4 orthologs (pg173)

(Adapted from Hughes et al., 2012)

Figure 8-3 Amino acid sequence alignment of the M- and C-motifs of human Dbf4 and Drf1 (pg174)

(Adapted from Hughes et al.).

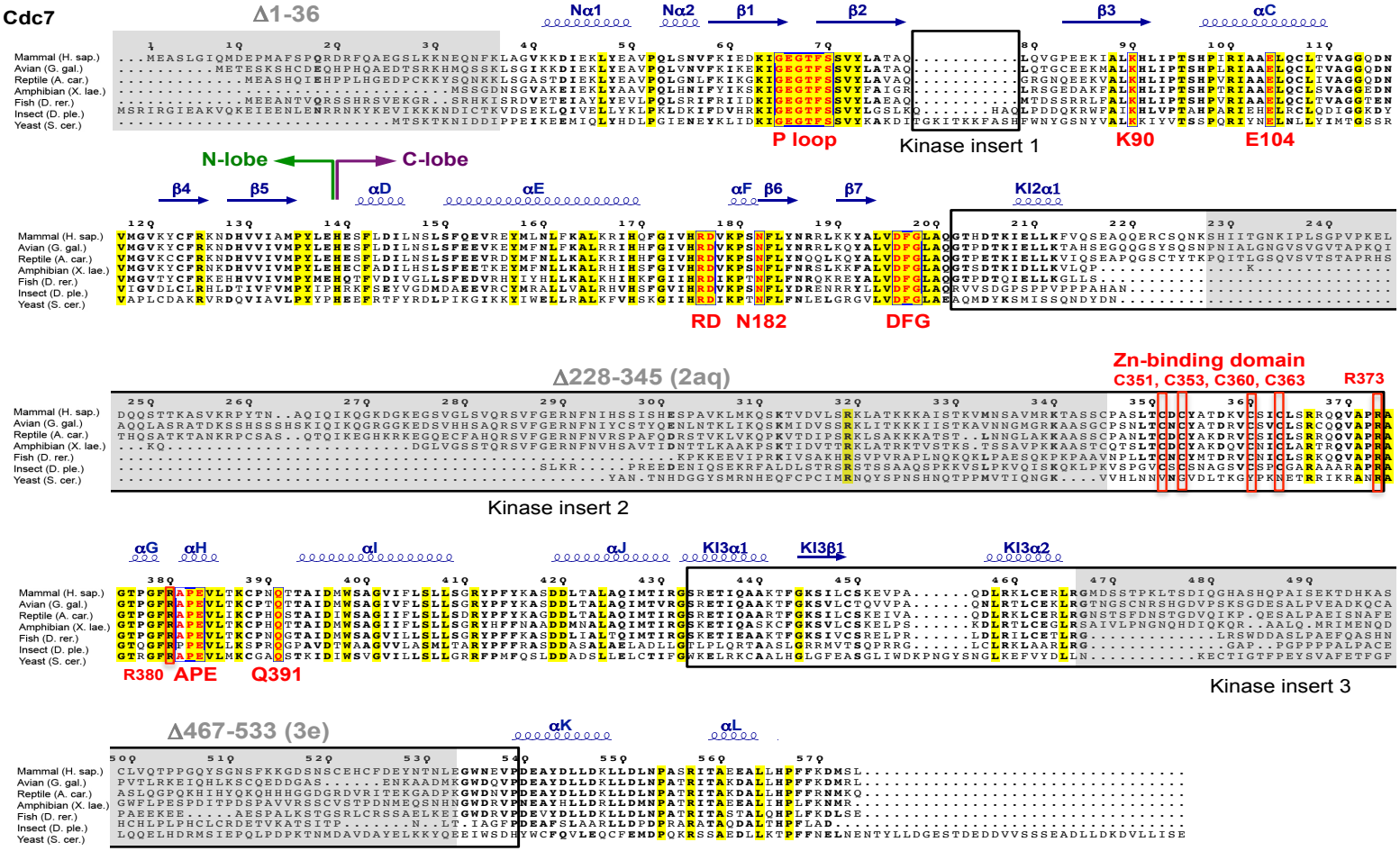
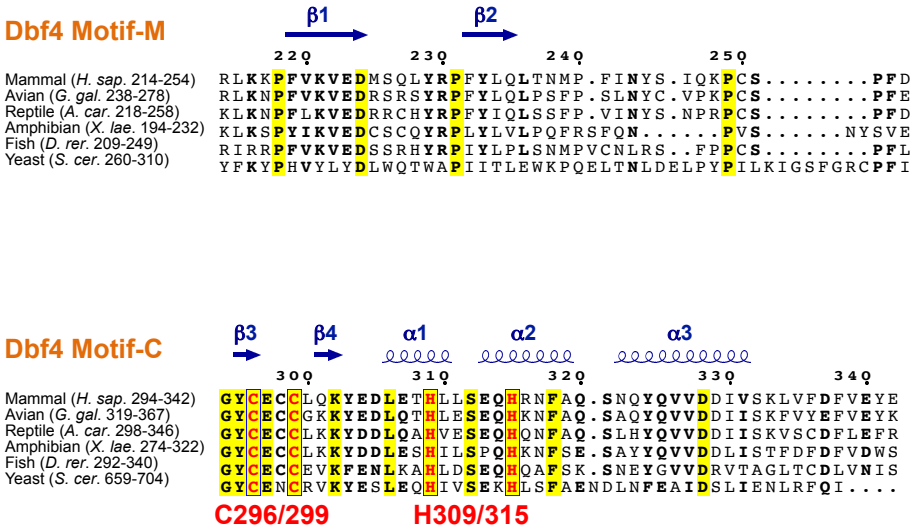


Figure 8-1 Amino acid sequence alignments of distal Cdc7 orthologues



8-2 Amino acid alignments of the motifs-M and C of distal orthologs Dbf4 orthologs (pg173)

Amino acid sequence alignment of the M- and C- motifs of human Dbf4 and Drf1

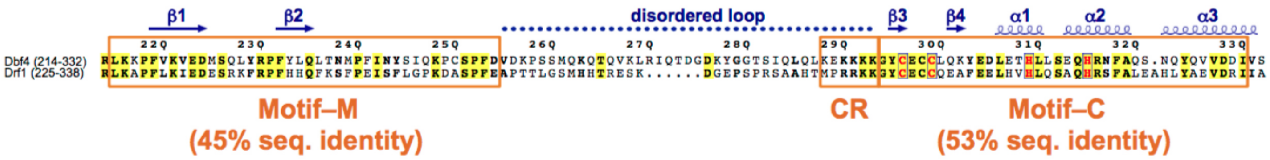


Figure 8-3 Amino acid sequence alignment of the M- and C-motifs of distal Dbf4 orthologues

Table 8-1 Table of plasmids used for protein expression in *E.coli.*

Plasmid Name	Use/Description	Reference
pRSFDuet1	Vector for protein co-expression in *E.coli*	Novagen
pCDFDuet1	Vector for protein co-expression in *E.coli*	Novagen
pRSF-Cdc7 (ΔN)	Expression of human Cdc7 (Δ1-36)	Hughes et al. 2012
pRSF-Cdc7 (ΔN/C345A)	Expression of human Cdc7 (Δ1-36/C345A)	This study
pRSF-Cdc7 (ΔN/C351A)	Expression of human Cdc7 (Δ1-36/C351A)	This study
pRSF-Cdc7 (ΔN/C353A)	Expression of human Cdc7 (Δ1-36/C353A)	This study
pRSF-Cdc7 (ΔN/C360A)	Expression of human Cdc7 (Δ1-36/C360A)	This study
pRSF-Cdc7 (ΔN/C363A)	Expression of human Cdc7 (Δ1-36/C363A)	This study
pRSF-Cdc7 (ΔN/R373A)	Expression of human Cdc7 (Δ1-36/R373A)	This study
pRSF-Cdc7 (ΔN/R380A)	Expression of human Cdc7 (Δ1-36/R380A)	This study
pRSF-Cdc7 (ΔN/R373A/R380A)	Expression of human Cdc7 (Δ1-36/C345A)	This study
pRSF-Cdc7 (ΔN/2n/3b)	Expression of human Cdc7 (Δ1-36/ Δ228-343/ Δ484-529)	Hughes et al. 2012
pRSF-Cdc7 (ΔN/2o/3b)	Expression of human Cdc7 (Δ1-36/ Δ228-349/ Δ484-529)	Hughes et al. 2012
pRSF-Cdc7 (ΔN/2q/3b)	Expression of human Cdc7 (Δ1-36/ Δ228-359/ Δ484-529)	Hughes et al. 2012
pRSF-Cdc7 (ΔN/2z/3b)	Expression of human Cdc7 (Δ1-36/ Δ228-305/ Δ484-529)	Hughes et al. 2012
pRSF-Cdc7 (ΔN/2ad/3b)	Expression of human Cdc7 (Δ1-36/ Δ228-282/ Δ484-529)	Hughes et al. 2012

pRSF-Cdc7 (ΔN/2ag/3b)	Expression of human Cdc7 (Δ1-36/ Δ228-337/ Δ484-529)	Hughes et al. 2012
pRSF-Cdc7 (ΔN/2ah/3b)	Expression of human Cdc7 (Δ1-36/ Δ228-275/ Δ484-529)	Hughes et al. 2012
pRSF-Cdc7 (ΔN/2ai/3b)	Expression of human Cdc7 (Δ1-36/ Δ228-285/ Δ484-529)	Hughes et al. 2012
pRSF-Cdc7 (ΔN/2ak/3b)	Expression of human Cdc7 (Δ1-36/ Δ228-295/ Δ484-529)	This study
pRSF-Cdc7 (ΔN/2al/3b)	Expression of human Cdc7 (Δ1-36/ Δ231-285/ Δ484-529)	This study
pRSF-Cdc7 (ΔN/2am/3b)	Expression of human Cdc7 (Δ1-36/ Δ231-382/ Δ484-529)	This study
pRSF-Cdc7 (ΔN/2an/3b)	Expression of human Cdc7 (Δ1-36/ Δ228-288/ Δ484-529)	This study
pRSF-Cdc7 (ΔN/2ai/3c)	Expression of human Cdc7 (Δ1-36/ Δ228-285/ Δ467-AGAG-532)	This study
pRSF-Cdc7 (ΔN/2o/3c)	Expression of human Cdc7 (Δ1-36/ Δ228-349/ Δ467-AGAG-532)	This study
pRSF-Cdc7 (ΔN/2y/3c)	Expression of human Cdc7 (Δ1-36/ Δ228-265/ Δ467-AGAG-532)	This study
pRSF-Cdc7 (ΔN/2z/3c)	Expression of human Cdc7 (Δ1-36/ Δ228-305/ Δ467-AGAG-532)	This study
pRSF-Cdc7 (ΔN/2o/3d)	Expression of human Cdc7 (Δ1-36/ Δ228-349/ Δ467-AGAGA-532)	This study
pRSF-Cdc7 (ΔN/2n/3e)	Expression of human Cdc7 (Δ1-36/ Δ228-343/ Δ467-AGAG-533)	This study
pRSF-Cdc7 (ΔN/2o/3e)	Expression of human Cdc7 (Δ1-36/ Δ228-349/ Δ467-AGAG-533)	This study
pRSF-Cdc7 (ΔN/2r/3e)	Expression of human Cdc7 (Δ1-36/ Δ234-343/ Δ467-AGAG-533)	This study
pRSF-Cdc7 (ΔN/2ao/3e)	Expression of human Cdc7 (Δ1-36/ Δ234-301/ Δ467-AGAG-533)	This study
pRSF-Cdc7 (ΔN/2ap/3e)	Expression of human Cdc7 (Δ1-36/ Δ234-326/ Δ467-AGAG-533)	This study
pRSF-Cdc7 (ΔN/2aq/3e)	Expression of human Cdc7 (Δ1-36/ Δ228-345/ Δ467-AGAG-533)	This study

pRSF-Cdc7 (ΔN/2ar/3e)	Expression of human Cdc7 (Δ1-36/ Δ228-347/ Δ467-AGAG-533)	This study
pRSF-Cdc7 (ΔN/2as/3e)	Expression of human Cdc7 (Δ1-36/ Δ232-349/ Δ467-AGAG-533)	This study
pRSF-Cdc7 (ΔN/2at/3e)	Expression of human Cdc7 (Δ1-36/ Δ230-349/ Δ467-AGAG-533)	This study
pRSF-Cdc7 (ΔN/2o/3f)	Expression of human Cdc7 (Δ1-36/ Δ228-349/ Δ466-AGA-533)	This study
pCDF-His-Dbf4 (MC)	Expression of human Dbf4 (210-350) with an N-terminal His_{6t}tag	Hughes et al. 2012
pCDF-His-Dbf4 (MC2)	Expression of human Dbf4 (214-344) with an N-terminal His_{6t}tag	This study
pCDF-His-Drf1 (141-348)	Expression of human Drf1 (141-348) with an N-terminal His_{6t}tag	This study
pCDF-His-Drf1 (165-348)	Expression of human Drf1 (165-348) with an N-terminal His_{6t}tag	This study
pCDF-His-Drf1 (178-348)	Expression of human Drf1 (178-348) with an N-terminal His_{6t}tag	This study
pCDF-His-Drf1 (211-348)	Expression of human Drf1 (211-348) with an N-terminal His_{6t}tag	This study
pCDF-His-Drf1 (211-345)	Expression of human Drf1 (211-345) with an N-terminal His_{6t}tag	This study
pCDF-His-Drf1 (214-345)	Expression of human Drf1 (214-345) with an N-terminal His_{6t}tag	This study
pCDF-His-Drf1 (214-348)	Expression of human Drf1 (214-348) with an N-terminal His_{6t}tag	This study

Table 8-2 Plasmids used for retrovirus and lentivirus production for expression in mammalian cells..

Plasmid Name	Use/Description	Reference
pWPT-GFP	Lentiviral vector expressing GFP	Addgene ♯1225 Didier Trono
pBABE (puro)	Retroviral vector	Morgensten et al. (1990)
pWZL (hygro)	Retroviral vector	Addgene plasmid #18750 Scott Lowe
pCG-GagPol	Retroviral packaging vector.	Ulm et al. 2007
p8.2	Lentiviral packaging vector	Naldini et al. 1996
pMD.G	Lentiviral envelope vector	Zufferey et al. 1997
pX459	Mammalian expression vector for CRISPR/Cas9	Addgene ♯62988 Feng Zhang
pQ-FLAG-Cdc7	Retroviral vector used for stable expression of human Cdc7 with an N-terminal FLAG tag	Hughes et al.
pWPT-mCherry-Cdc7	Lentiviral vector for stable expression of human Cdc7 with an N-terminal mCherry fusion.	This study

pBABE-FLAG-Cdc7 (WT)	Retroviral vector for stable expression of human Cdc7 with an N-terminal FLAG tag	This study
pBABE-FLAG-Cdc7 (D196N)	Retroviral vector for stable expression of human Cdc7 (D196N) with an N-terminal FLAG tag	This study
pBABE-FLAG-Cdc7 (C353A)	Retroviral vector for stable expression of human Cdc7 (C353A) with an N-terminal FLAG tag	This study
pBABE-FLAG-Cdc7 (C360A)	Retroviral vector for stable expression of human Cdc7 (C360A) with an N-terminal FLAG tag	This study
pBABE-FLAG-Cdc7 (R380A)	Retroviral vector for stable expression of human Cdc7 (R380A) with an N-terminal FLAG tag	This study
pBABE-FLAG-Cdc7 (R373A/R380A)	Retroviral vector for stable expression of human Cdc7 (R373A/R380A) with an N-terminal FLAG tag	This study
pBABE-FLAG-Cdc7 (Δ2aq)	Retroviral vector for stable expression of human Cdc7 (Δ2aq) with an N-terminal FLAG tag	This study
pBABE-FLAG-Cdc7 (Δ2aq/3e)	Retroviral vector for stable expression of human Cdc7 (Δ2aq/3e) with an N-terminal FLAG tag	This study
pWZL-FLAG-Cdc7 (Δ2aq)	Retroviral vector for stable expression of human Cdc7 (Δ2aq) with an N-terminal FLAG tag	This study

pWZL-FLAG-Cdc7 (Δ2aq/3e)	Retroviral vector for stable expression of human Cdc7 (Δ2aq/3e) with an N-terminal FLAG tag	This study
pBABE-FLAG-Cdc7 (Δ2aq/Kz1)	Retroviral vector for stable expression of human Cdc7 (Δ2aq) with an N-terminal FLAG tag and a single point mutation in the Kozak consensus sequence.	This study
pBABE-FLAG-Cdc7 (Δ2aq/Kz2)	Retroviral vector for stable expression of human Cdc7 (Δ2aq) with an N-terminal FLAG tag and a double point mutation in the Kozak consensus sequence.	This study
pBABE-FLAG-Cdc7 (Δ2aq/3e/Kz1)	Retroviral vector for stable expression of human Cdc7 (Δ2aq/3e) with an N-terminal FLAG tag and a single point mutation in the Kozak consensus sequence.	This study
pBABE-FLAG-Cdc7 (Δ2aq/3e/Kz2)	Retroviral vector for stable expression of human Cdc7 (Δ2aq/3e) with an N-terminal FLAG tag and a double point mutation in the Kozak consensus sequence.	This study

Table 8-3 Cell lines produced and used as part of this study.

Cell Line Name	Description	Reference
HT1080 mCherry-Cdc7(LoxP)	HT1080 cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site.	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site.	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7.	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (D196N)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (D196N)	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (C353A)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (C353A).	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (R373A)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (R373A).	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (R373A/R380A)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (R373A/R380A).	This study

HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (Δ2aq) (pBABE)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (Δ2aq).	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (Δ2aq/3e) (pBABE)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (Δ2aq/3e).	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (Δ2aq) (pBABE/Kz1)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (Δ2aq) with a single point mutation in the Kozak consensus sequence.	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (Δ2aq) (pBABE/Kz2)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (Δ2aq) with a double point mutation in the Kozak consensus sequence.	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (Δ2aq/3e) (pBABE/Kz1)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (Δ2aq/3e) with a single point mutation in the Kozak consensus sequence.	This study

HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (Δ2aq/3e) (pBABE/Kz2)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (Δ2aq/3e) with a double point mutation in the Kozak consensus sequence.	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (Δ2aq) (pWZL)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (Δ2aq)	This study
HT1080 *CDC7*(-/-) mCherry-Cdc7(LoxP) + FLAG-Cdc7 (Δ2aq/3e) (pWZL)	HT1080 (Cdc7-/-) cells stably expressing human Cdc7 with an N-terminal mCherry fusion with LoxP site and FLAG-Cdc7 (Δ2aq/3e)	This study

Table 8-4 Oligos used in this study.

Oligo	Sequence 5'-3'	Description	Restriction site	Reference
SH78	GCCAGCCCATGGGACTTGCAGGTGTTAAAAAAGATATTG	pRSF-Cdc7 (Δ1-36) construction	NcoI	Hughes et al. (2012)
SH81	GGAGCCTCGAGTCACAAGCTCATATCTTTAAAAAATGGATGC	pRSF-Cdc7 (C-terminus)	XhoI	Hughes et al. (2012)
SD1	GAAAACTGCCAGTTCTGCCCCAGCTAGCCTGAC	Cdc7 C345A mutation	n/a	This study
SD2	GTCAGGCTAGCTGGGGCAGAACTGGCAGTTTTC	Cdc7 C345A mutation	n/a	This study
SD3	CAGCTAGCCTGACCGCTGACTGCTATGCAAC	Cdc7 C351A mutation	n/a	This study
SD4	GTTGCATAGCAGTCAGCGGTCAGGCTAGCTG	Cdc7 C351A mutation	n/a	This study
SD5	CTAGCCTGACCTGTGACGCCTATGCAACAGATAAAG	Cdc7 C353A mutation	n/a	This study
SD6	CTTTATCTGTTGCATAGGCGTCACAGGTCAGGCTAG	Cdc7 C353A mutation	n/a	This study
SD11	CTATGCAACAGATAAAGTTGCCAGTATTTGCCTTTCAAGGCG	Cdc7 C360A mutation	n/a	This study
SD12	CGCCTTGAAAGGCAAATACTGGCAACTTTATCTGTTGCATAG	Cdc7 C360A mutation	n/a	This study
SD13	GATAAAGTTTGTAGTATTGCCCTTTCAAGGCGTCAGC	Cdc7 C363A mutation	n/a	This study
SD14	GCTGACGCCTTGAAAGGGCAATACTACAAACTTTATC	Cdc7 C363A mutation	n/a	This study
SD84	TCAGCAGGTTGCCCCTGCAGCAGGTACACCAGGAT	Cdc7 R373A mutation	n/a	This study

SD85	ATCCTGGTGTACCTGCTGCAGGGGCAACCTGCTGA	Cdc7 R373A mutation	n/a	This study
SD86	GTACACCAGGATTCGCAGCACCAGAGGTC	Cdc7 R380A mutation	n/a	This study
SD87	GACCTCTGGTGCTGCGAATCCTGGTGTAC	Cdc7 R380A mutation	n/a	This study
SH119	GGTGTTCACAAAACAAATCTTGCCCAGCTAGCCTG	Cdc7 2n deletion	n/a	This study
SH120	CAGGCTAGCTGGGCAAGATTTGTTTTGTGAACACC	Cdc7 2n deletion	n/a	This study
SH121	GGTGTTCACAAAACAAAACCTGTGACTGCTATGC	Cdc7 2o deletion	n/a	This study
SH122	GCATAGCAGTCACAGGTTTTGTTTTGTGAACACC	Cdc7 2o deletion	n/a	This study
SH125	GGTGTTCACAAAACAAATGTAGTATTTGCCTTTCAAG	Cdc7 2q deletion	n/a	This study
SH126	CTTGAAAGGCAAATACTACATTTGTTTTGTGAACACC	Cdc7 2q deletion	n/a	This study
SH141	GGTGTTCACAAAACAAACAGATTAAACAAGGAAAAG	Cdc7 2y deletion	n/a	This study
SH142	CTTTTCCTTGTTTAATCTGTTTGTTTTGTGAACACC	Cdc7 2y deletion	n/a	This study
SH127	CCCACATAATCACAGGATCTTGCCCAGCTAGCCTG	Cdc7 2r deletion	n/a	This study
SH128	CAGGCTAGCTGGGCAAGATCCTGTGATTATGTGGG	Cdc7 2r deletion	n/a	This study
SH159	GGTGTTCACAAAACAAAAAACTCATGAAGCAGTC	Cdc7 2z deletion	n/a	This study
SH160	GACTGCTTCATGAGTTTTTTGTTTTGTGAACACC	Cdc7 2z deletion	n/a	This study
SH197	GGTGTTCACAAAACAAACAGCGCTCTGTTTTTGG	Cdc7 2ad deletion	n/a	This study
SH198	CCAAAAACAGAGCGCTGTTTGTTTTGTGAACACC	Cdc7 2ad deletion	n/a	This study
SH203	GGTGTTCACAAAACAAAATGAGGAAAACTGCCAG	Cdc7 2ag deletion	n/a	This study

SH204	CTGGCAGTTTTCCTCATTTTGTTTTGTGAACACC	Cdc7 2ag deletion	n/a	This study
SH205	GGTGTTCACAAAACAAAGGATCTGTAGGCCTTTCTG	Cdc7 2ah deletion	n/a	This study
SH206	CAGAAAGGCCTACAGATCCTTTGTTTTGTGAACACC	Cdc7 2ah deletion	n/a	This study
SH207	GGTGTTCACAAAACAAAGTTTTTGGAGAAAGAAATTTC	Cdc7 2ai deletion	n/a	This study
SH208	GAAATTTCTTTCTCCAAAAACTTTGTTTTGTGAACACC	Cdc7 2ai deletion	n/a	This study
SH209	GGTGTTCACAAAACAAAAGCTCCATTTCACATGAG	Cdc7 2ak deletion	n/a	This study
SH210	CTCATGTGAAATGGAGCTTTTGTTTTGTGAACACC	Cdc7 2ak deletion	n/a	This study
PC795	GTGTTCACAAAACAAATCCCACATAGTTTTTGGAGAAAGAAATTTCAATATAC	Cdc7 2al deletion	n/a	This study
PC796	GTATATTGAAATTTCTTTCTCCAAAAACTATGTGGGATTTGTTTTGTGAACAC	Cdc7 2al deletion	n/a	This study
PC797	GTGTTCACAAAACAAATCCCACATACAGCGCTCTGTTTTTGGA	Cdc7 2am deletion	n/a	This study
PC798	GTGTTCACAAAACAAATCCCACATACAGCGCTCTGTTTTTGGA	Cdc7 2am deletion	n/a	This study
PC799	GGTGTTCACAAAACAAAGAAAGAAATTTCAATATACAC	Cdc7 2an deletion	n/a	This study
PC800	GTGTATATTGAAATTTCTTTCTTTGTTTTGTGAACACC	Cdc7 2an deletion	n/a	This study
SD34	CACATAATCACAGGAAACAGCCCTGCAGTGAAAC	Cdc7 2ao deletion	n/a	This study
SD35	GTTTCACTGCAGGGCTGTTTCCTGTGATTATGTG	Cdc7 2ao deletion	n/a	This study
SD36	CACATAATCACAGGAAACGCTATTTCTACAAAAG	Cdc7 2ap deletion	n/a	This study

SD37	CTTTTGTAGAAATAGCGTTTCCTGTGATTATGTG	Cdc7 2ap deletion	n/a	This study
SD45	GGTGTTCACAAAACAAACCAGCTAGCCTGACCTGTGACTGCT	Cdc7 2aq deletion	n/a	This study
SD46	AGCAGTCACAGGTCAGGCTAGCTGGTTTGTTTTGTGAACACC	Cdc7 2aq deletion	n/a	This study
SD47	GGTGTTCACAAAACAAAAGCCTGACCTGTGACTGCTATGC	Cdc7 2ar deletion	n/a	This study
SD48	GCATAGCAGTCACAGGTCAGGCTTTTGTTTTGTGAACACC	Cdc7 2ar deletion	n/a	This study
SD49	TGTTCACAAAACAAATCCCACATAATCACCTGTGACTGCTATGCA	Cdc7 2as deletion	n/a	This study
SD50	TGCATAGCAGTCACAGGTGATTATGTGGGATTTGTTTTGTGAAC	Cdc7 2as deletion	n/a	This study
SD51	GTGTTCACAAAACAAATCCCACACCTGTGACTGCTATGCA	Cdc7 2at deletion	n/a	This study
SD52	TGCATAGCAGTCACAGGTGTGGGATTTGTTTTGTGAACAC	Cdc7 2at deletion	n/a	This study
SH62	GATATACAAGGGCATGCTACCAATTTAGAAGGCTGG	Cdc7 3b deletion	n/a	This study
SH63	CCAGCCTTCTAAATTGGTAGCATGCCCTTGTATATC	Cdc7 3b deletion	n/a	This study
SD7	TGTGAGAGACTCAGGGGTGCAGGTGCAGGTGAAGGCTGGAATGAGGTA	Cdc7 3c deletion	n/a	This study
SD8	TACCTCATTCCAGCCTTCACCTGCACCTGCACCCCTGAGTCTCTCACA	Cdc7 3c deletion	n/a	This study
SD9	GAGAGACTCAGGGGTGCAGGTGCAGGTGCAGAAGGCTGGAATGAGGTA	Cdc7 3d deletion	n/a	This study

SD10	TACCTCATTCCAGCCTTCTGCACCTGCACCTGCACCCCTGAGTC TCTC	Cdc7 3d deletion	n/a	This study
SD21	GAGAGACTCAGGGGTGCAGGTGCAGGTGGCTGGAATGAGGTA	Cdc7 3e deletion	n/a	This study
SD22	TACCTCATTCCAGCCACCTGCACCTGCACCCCTGAGTCTCT	Cdc7 3e deletion	n/a	This study
SD23	TGTGAGAGACTCAGGGCAGGTGCAGGCTGGAATGAGGTACCT	Cdc7 3f deletion	n/a	This study
SD24	AGGTACCTCATTCCAGCCTGCACCTGCCCTGAGTCTCTCACA	Cdc7 3f deletion	n/a	This study
SH178	GGAC CCCGGG ACAAGAACAGGAAGACTCAAAAAGCC	pCDF-His-Dbf4 210->	XmaI/SmaI	This study
SH44	GGTC CTCGAG CTATCTTTTCTTTTTAGGTGTGTCCTTTTC	pCDF-His-Dbf4 ->350	XhoI	This study
SD25	GGACCCCGGGAGACTCAAAAAGCCTTTTGTAAAGG	pCDF-His-Dbf4 214->	XmaI/SmaI	This study
SD20	GGTCCTCGAGCTAGTCCTTTTCATATTCCACAAAGTC	pCDF-His-Dbf4 ->344	XhoI	This study
SD109	GGACCCCGGGCCTCTAAGCAGAGGGAAGGAGCTGCTGCA	pCDF-His-Drf1 141->	XmaI/SmaI	This study
SD165	GGACCCCGGGGGGGGCAGCAGCAGCCTCCTGACCAATG	pCDF-His-Drf1 165->	XmaI/SmaI	This study
SD141	GGACCCCGGGGGAGTGAGGATTCTGCACGT	pCDF-His-Drf1 178->	XmaI/SmaI	This study
SD107	GGACCCCGGGGGAACATGTCCAGCAGCAGAGTCAAGAAC	pCDF-His-Drf1 211->	XmaI/SmaI	This study
SD142	GGACCCCGGGCCAGCAGCAGAGTCAAGAACACGGA	pCDF-His-Drf1 214->	XmaI/SmaI	This study
SD143	GGTCCTCGAGTGCAAAGCTGTGGCTGAGCTGAGCA	pCDF-His-Drf1 ->345	XhoI	This study
SD110	GGTCCTCGAGAGGGATGTCTGCAAAGCTGTGGCT	pCDF-His-Drf1 ->348	XhoI	This study
SD38	GGCGACGCGTCGCCACCATGGTGAGCAAGGGCGAGGAGGATA	pWPT-mCherry-Cdc7 N-	MluI	This study

		terminus		
SD16	TGAATACCCAGGCTTGCTTCCTTGTACAGCTCGTCCATGCCGCC	mCherry C-terminus with codon optimised Cdc7 overhang for fusion PCR	n/a	This study
SD17	ATGGACGAGCTGTACAAGGAAGCAAGCCTGGGTATTCAGA	Codon optimised Cdc7 N-terminus with mCherry overhang for Fusion PCR	n/a	This study
SD31	GGCGGTCGACTTACAGGCTCATATCTTTGAAAAACGG	pWPT-mCherry-Cdc7 C-terminus	SalI	This study
SD105	GAC TACGTA CGCACCATGGATTACAAGGATGACG	pBABE-Flag-Cdc7 N-terminus	SnaBI	This Study
SD96	GGAC GTCGAC TCACAAGCTCATATCTTTAAAAA	pBABE-FLAG-Cdc7 C-terminus	SalI	This study
SD150	GAC TACGTA ATGGAGGCGTCTTTGGGG	pWZL-FLAG-Cdc7 N-terminus	SnaBI	This Study
SD147	CTCACAAGCTCATATCTTTAAAAAATGGATG	pWZL-FLAG-Cdc7 N-terminus	BamHI	This Study

Table 8-5 Peptides used in *in vitro* kinase assays

Peptide Name	Residues	Modifications	Use
Mcm2-S40-assay	Mcm2 35-47	N-terminal biotin, S41p	Kinase assays for crystallography constructs (Hughes et al. 2012)
Mcm2-S40(15)	Mcm2 33-47	N-terminal biotin, S41p	Kinase assays for point mutants of the P+4 residue (positive control)
Mcm2-S40(15)NoP	Mcm2 33-47	N-terminal biotin	Kinase assays for point mutants of the P+4 residue (no pre-phosphorylation of S41)
Mcm2-S40(15) R44A	Mcm2 33-47	N-terminal biotin, S41p, R44A	Kinase assays for point mutants of the P+4 residue
Mcm2-S40(15) R44E	Mcm2 33-47	N-terminal biotin, S41p, R44E	Kinase assays for point mutants of the P+4 residue
Mcm2-S40(15) R44K	Mcm2 33-47	N-terminal biotin, S41p, R44K	Kinase assays for point mutants of the P+4 residue
Mcm2-S40(15) R44L	Mcm2 33-47	N-terminal biotin, S41p, R44L	Kinase assays for point mutants of the P+4 residue
Mcm2-S40(15) R44F	Mcm2 33-47	N-terminal biotin, S41p, R44F	Kinase assays for point mutants of the P+4 residue

Reference List

ABID ALI, F., RENAULT, L., GANNON, J., GAHLON, H. L., KOTECHA, A., ZHOU, J. C., RUEDA, D. & COSTA, A. 2016. Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. *Nat Commun,* 7**,** 10708.

ADAMS, J. A. 2001. Kinetic and catalytic mechanisms of protein kinases. *Chem Rev,* 101**,** 2271-90.

ADAMS, J. A. 2003. Activation loop phosphorylation and catalysis in protein kinases: is there functional evidence for the autoinhibitor model? *Biochemistry,* 42**,** 601-7.

ADAMS, P. D., AFONINE, P. V., BUNKOCZI, G., CHEN, V. B., DAVIS, I. W., ECHOLS, N., HEADD, J. J., HUNG, L.-W., KAPRAL, G. J., GROSSE-KUNSTLEVE, R. W., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R., READ, R. J., RICHARDSON, D. C., RICHARDSON, J. S., TERWILLIGER, T. C. & ZWART, P. H. 2010a. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D,* 66**,** 213-221.

ADAMS, P. D., AFONINE, P. V., BUNKOCZI, G., CHEN, V. B., DAVIS, I. W., ECHOLS, N., HEADD, J. J., HUNG, L. W., KAPRAL, G. J., GROSSE-KUNSTLEVE, R. W., MCCOY, A. J., MORIARTY, N. W., OEFFNER, R., READ, R. J., RICHARDSON, D. C., RICHARDSON, J. S., TERWILLIGER, T. C. & ZWART, P. H. 2010b. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D-Biological Crystallography,* 66**,** 213-221.

AGGARWAL, B. D. & CALVI, B. R. 2004. Chromatin regulates origin activity in Drosophila follicle cells. *Nature,* 430**,** 372-376.

AMMOSOVA, T., OBUKHOV, Y., KOTELKIN, A., BREUER, D., BEULLENS, M., GORDEUK, V. R., BOLLEN, M. & NEKHAI, S. 2011. Protein phosphatase-1 activates CDK9 by dephosphorylating Ser175. *PLoS One,* 6**,** e18985.

ANTON, M. & GRAHAM, F. L. 1995. Site-specific recombination mediated by an adenovirus vector expressing the Cre recombinase protein: a molecular switch for control of gene expression. *J Virol,* 69**,** 4600-6.

ARIAS, E. E. & WALTER, J. C. 2006. PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. *Nature Cell Biology,* 8**,** 84-U33.

ARIAS, E. E. & WALTER, J. C. 2007. Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev,* 21**,** 497-518.

BAHMAN, M., BUCK, V., WHITE, A. & ROSAMOND, J. 1988. Characterization of the Cdc7 Gene-Product of Saccharomyces-Cerevisiae as a Protein-Kinase Needed for the Initiation of Mitotic DNA-Synthesis. *Biochimica Et Biophysica Acta,* 951**,** 335-343.

BAILIS, J. M., BERNARD, P., ANTONELLI, R., ALLSHIRE, R. C. & FORSBURG, S. L. 2003a. Hsk1-Dfp1 is required for heterochromatin-mediated cohesion at centromeres. *Nature Cell Biology,* 5**,** 1111-1116.

BAILIS, J. M., BERNARD, P., ANTONELLI, R., ALLSHIRE, R. C. & FORSBURG, S. L. 2003b. Hsk1-Dfp1 is required for heterochromatin-mediated cohesion at centromeres. *Nat Cell Biol,* 5**,** 1111-6.

BALESTRINI, A., COSENTINO, C., ERRICO, A., GARNER, E. & COSTANZO, V. 2010. GEMC1 is a TopBP1-interacting protein required for chromosomal DNA replication. *Nat Cell Biol,* 12**,** 484-91.

BAO, Z. Q., JACOBSEN, D. M. & YOUNG, M. A. 2011. Briefly bound to activate: transient binding of a second catalytic magnesium activates the structure and dynamics of CDK2 kinase for catalysis. *Structure,* 19**,** 675-90.

BARALDI, E., DJINOVIC CARUGO, K., HYVONEN, M., SURDO, P. L., RILEY, A. M., POTTER, B. V., O'BRIEN, R., LADBURY, J. E. & SARASTE, M. 1999. Structure of the PH domain from Bruton's tyrosine kinase in complex with inositol 1,3,4,5-tetrakisphosphate. *Structure,* 7**,** 449-60.

BARIK, S. 1993. Expression and biochemical properties of a protein serine/threonine phosphatase encoded by bacteriophage lambda. *Proc Natl Acad Sci U S A,* 90**,** 10633-7.

BARRANGOU, R., FREMAUX, C., DEVEAU, H., RICHARDS, M., BOYAVAL, P., MOINEAU, S., ROMERO, D. A. & HORVATH, P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science,* 315**,** 1709-12.

BAYLISS, R., SARDON, T., VERNOS, I. & CONTI, E. 2003. Structural basis of Aurora-A activation by TPX2 at the mitotic spindle. *Molecular Cell,* 12**,** 851-862.

BELL, S. P. & STILLMAN, B. 1992. Atp-Dependent Recognition of Eukaryotic Origins of DNA-Replication by a Multiprotein Complex. *Nature,* 357**,** 128-134.

BELLON, S., FITZGIBBON, M. J., FOX, T., HSIAO, H. M. & WILSON, K. P. 1999. The structure of phosphorylated p38gamma is monomeric and reveals a conserved activation-loop conformation. *Structure,* 7**,** 1057-65.

BERNARD, P., MAURE, J. F., PARTRIDGE, J. F., GENIER, S., JAVERZAT, J. P. & ALLSHIRE, R. C. 2001. Requirement of heterochromatin for cohesion at centromeres. *Science,* 294**,** 2539-42.

BERRY, M. B. & PHILLIPS, G. N., JR. 1998. Crystal structures of Bacillus stearothermophilus adenylate kinase with bound Ap5A, Mg2+ Ap5A, and Mn2+ Ap5A reveal an intermediate lid position and six coordinate octahedral geometry for bound Mg2+ and Mn2+. *Proteins,* 32**,** 276-88.

BIONDI, R. M., CHEUNG, P. C., CASAMAYOR, A., DEAK, M., CURRIE, R. A. & ALESSI, D. R. 2000. Identification of a pocket in the PDK1 kinase domain that interacts with PIF and the C-terminal residues of PKA. *EMBO J,* 19**,** 979-88.

BIONDI, R. M. & NEBREDA, A. R. 2003. Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. *Biochem J,* 372**,** 1-13.

BLAKE-HODEK, K. A., WILLIAMS, B. C., ZHAO, Y., CASTILHO, P. V., CHEN, W., MAO, Y., YAMAMOTO, T. M. & GOLDBERG, M. L. 2012. Determinants for activation of the atypical AGC kinase Greatwall during M phase entry. *Mol Cell Biol,* 32**,** 1337-53.

BLOW, J. J. 1993. Preventing re-replication of DNA in a single cell cycle: evidence for a replication licensing factor. *J Cell Biol,* 122**,** 993-1002.

BLOW, J. J. & GILLESPIE, P. J. 2008. Replication licensing and cancer--a fatal entanglement? *Nat Rev Cancer,* 8**,** 799-806.

BONTE, D., LINDVALL, C., LIU, H., DYKEMA, K., FURGE, K. & WEINREICH, M. 2008. Cdc7-Dbf4 kinase overexpression in multiple cancers and tumor cell lines is correlated with p53 inactivation. *Neoplasia,* 10**,** 920-31.

BOOS, D., FRIGOLA, J. & DIFFLEY, J. F. X. 2012. Activation of the replicative DNA helicase: breaking up is hard to do. *Current Opinion in Cell Biology,* 24**,** 423-430.

BOOS, D., SANCHEZ-PULIDO, L., RAPPAS, M., PEARL, L. H., OLIVER, A. W., PONTING, C. P. & DIFFLEY, J. F. X. 2011. Regulation of DNA Replication through Sld3-Dpb11 Interaction Is Conserved from Yeast to Humans. *Current Biology,* 21**,** 1152-1157.

BOUSSET, K. & DIFFLEY, J. F. X. 1998. The Cdc7 protein kinase is required for origin firing during S phase. *Genes & Development,* 12**,** 480-490.

BRAMAN, J., PAPWORTH, C. & GREENER, A. 1996. Site-directed mutagenesis using double-stranded plasmid DNA templates. *Methods Mol Biol,* 57**,** 31-44.

BREWER, B. J. & FANGMAN, W. L. 1987. The Localization of Replication Origins on Ars Plasmids in Saccharomyces-Cerevisiae. *Cell,* 51**,** 463-471.

BROWN, G. W. & KELLY, T. J. 1999. Cell cycle regulation of Dfp1, an activator of the Hsk1 protein kinase. *Proc Natl Acad Sci U S A,* 96**,** 8443-8.

BROWN, N. R., NOBLE, M. E., ENDICOTT, J. A. & JOHNSON, L. N. 1999a. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat Cell Biol,* 1**,** 438-43.

BROWN, N. R., NOBLE, M. E. M., ENDICOTT, J. A. & JOHNSON, L. N. 1999b. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nature Cell Biology,* 1**,** 438-443.

BUCK, V., WHITE, A. & ROSAMOND, J. 1991a. CDC7 protein kinase activity is required for mitosis and meiosis in Saccharomyces cerevisiae. *Mol Gen Genet,* 227**,** 452-7.

BUCK, V., WHITE, A. & ROSAMOND, J. 1991b. Cdc7 Protein-Kinase Activity Is Required for Mitosis and Meiosis in Saccharomyces-Cerevisiae. *Molecular & General Genetics,* 227**,** 452-457.

BURGESS, A., VIGNERON, S., BRIOUDES, E., LABBE, J. C., LORCA, T. & CASTRO, A. 2010. Loss of human Greatwall results in G2 arrest and multiple mitotic defects due to deregulation of the cyclin B-Cdc2/PP2A balance. *Proc Natl Acad Sci U S A,* 107**,** 12564-9.

CANAGARAJAH, B. J., KHOKHLATCHEV, A., COBB, M. H. & GOLDSMITH, E. J. 1997. Activation mechanism of the MAP kinase ERK2 by dual phosphorylation. *Cell,* 90**,** 859-869.

CHAN, K. F., HURST, M. O. & GRAVES, D. J. 1982. Phosphorylase kinase specificity. A comparative study with cAMP-dependent protein kinase on synthetic peptides and peptide analogs of glycogen synthase and phosphorylase. *J Biol Chem,* 257**,** 3655-9.

CHANG, F. J., RIERA, A., EVRIN, C., SUN, J. C., LI, H. L., SPECK, C. & WEINREICH, M. 2015. Cdc6 ATPase activity disengages Cdc6 from the pre-replicative complex to promote DNA replication. *Elife,* 4.

CHANTALAT, L., LEROY, D., FILHOL, O., NUEDA, A., BENITEZ, M. J., CHAMBAZ, E. M., COCHET, C. & DIDEBERG, O. 1999. Crystal structure of the human protein kinase CK2 regulatory subunit reveals its zinc finger-mediated dimerization. *EMBO J,* 18**,** 2930-40.

CHARYCH, D. H., COYNE, M., YABANNAVAR, A., NARBERES, J., CHOW, S., WALLROTH, M., SHAFER, C. & WALTER, A. O. 2008. Inhibition of Cdc7/Dbf4 kinase activity affects specific phosphorylation sites on MCM2 in cancer cells. *J Cell Biochem,* 104**,** 1075-86.

CHEN, S. Y. & BELL, S. P. 2011. CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6. *Genes & Development,* 25**,** 363-372.

CHEN, V. B., ARENDALL, W. B., HEADD, J. J., KEEDY, D. A., IMMORMINO, R. M., KAPRAL, G. J., MURRAY, L. W., RICHARDSON, J. S. & RICHARDSON, D. C. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D-Biological Crystallography,* 66**,** 12-21.

CHEN, Y. C. & WEINREICH, M. 2010. Dbf4 regulates the Cdc5 Polo-like kinase through a distinct non-canonical binding interaction. *J Biol Chem,* 285**,** 41244-54.

CHENG, A. N., JIANG, S. S., FAN, C. C., LO, Y. K., KUO, C. Y., CHEN, C. H., LIU, Y. L., LEE, C. C., CHEN, W. S., HUANG, T. S., WANG, T. Y. & LEE, A. Y. 2013. Increased Cdc7 expression is a marker of oral squamous cell carcinoma and overexpression of Cdc7 contributes to the resistance to DNA-damaging agents. *Cancer Lett,* 337**,** 218-25.

CHINI, C. C. & CHEN, J. 2003. Human claspin is required for replication checkpoint control. *J Biol Chem,* 278**,** 30057-62.

CHINI, C. C. & CHEN, J. 2004. Claspin, a regulator of Chk1 in DNA replication stress pathway. *DNA Repair (Amst),* 3**,** 1033-7.

CHO, W. H., LEE, Y. J., KONG, S. I., HURWITZ, J. & LEE, J. K. 2006. CDC7 kinase phosphorylates serine residues adjacent to acidic amino acids in the minichromosome maintenance 2 protein. *Proc Natl Acad Sci U S A,* 103**,** 11521-6.

CHOU, F. L., HILL, J. M., HSIEH, J. C., POUYSSEGUR, J., BRUNET, A., GLADING, A., UBERALL, F., RAMOS, J. W., WERNER, M. H. & GINSBERG, M. H. 2003. PEA-15 binding to ERK1/2 MAPKs is required for its modulation of integrin activation. *Journal of Biological Chemistry,* 278**,** 52587-52597.

CHOWDHURY, A., LIU, G., KEMP, M., CHEN, X., KATRANGI, N., MYERS, S., GHOSH, M., YAO, J., GAO, Y., BUBULYA, P. & LEFFAK, M. 2010. The DNA unwinding element binding protein DUE-B interacts with Cdc45 in preinitiation complex formation. *Mol Cell Biol,* 30**,** 1495-507.

CHUANG, L. C., TEIXEIRA, L. K., WOHLSCHLEGEL, J. A., HENZE, M., YATES, J. R., MENDEZ, J. & REED, S. I. 2009. Phosphorylation of Mcm2 by Cdc7 promotes pre-replication complex assembly during cell-cycle re-entry. *Mol Cell,* 35**,** 206-16.

CLARK-LEWIS, I., SANGHERA, J. S. & PELECH, S. L. 1991. Definition of a consensus sequence for peptide substrate recognition by p44mpk, the meiosis-activated myelin basic protein kinase. *J Biol Chem,* 266**,** 15180-4.

CLARKE, L. E., FOUNTAINE, T. J., HENNESSY, J., BRUGGEMAN, R. D., CLARKE, J. T., MAUGER, D. T. & HELM, K. F. 2009. Cdc7 expression in melanomas, Spitz tumors and melanocytic nevi. *J Cutan Pathol,* 36**,** 433-8.

COHEN, P. 2002. The origins of protein phosphorylation. *Nat Cell Biol,* 4**,** E127-30.

COSTANZO, M., NISHIKAWA, J. L., TANG, X. L., MILLMAN, J. S., SCHUB, O., BREITKREUZ, K., DEWAR, D., RUPES, I., ANDREWS, B. & TYERS, M. 2004. CDK activity antagonizes Whi5, an inhibitor of G1/S transcription in yeast. *Cell,* 117**,** 899-913.

COSTANZO, V., SHECHTER, D., LUPARDUS, P. J., CIMPRICH, K. A., GOTTESMAN, M. & GAUTIER, J. 2003. An ATR- and Cdc7-dependent DNA damage checkpoint that inhibits initiation of DNA replication. *Mol Cell,* 11**,** 203-13.

COSTER, G., FRIGOLA, J., BEURON, F., MORRIS, E. P. & DIFFLEY, J. F. X. 2014. Origin Licensing Requires ATP Binding and Hydrolysis by the MCM Replicative Helicase. *Molecular Cell,* 55**,** 666-677.

CRAIG, A. W., COSENTINO, G. P., DONZE, O. & SONENBERG, N. 1996. The kinase insert domain of interferon-induced protein kinase PKR is required for activity but not for interaction with the pseudosubstrate K3L. *J Biol Chem,* 271**,** 24526-33.

D'ARCY, A., BERGFORS, T., COWAN-JACOB, S. W. & MARSH, M. 2014. Microseed matrix screening for optimization in protein crystallization: what have we learned? *Acta Crystallogr F Struct Biol Commun,* 70**,** 1117-26.

DAVE, A., COOLEY, C., GARG, M. & BIANCHI, A. 2014. Protein phosphatase 1 recruitment by Rif1 regulates DNA replication origin firing by counteracting DDK activity. *Cell Rep,* 7**,** 53-61.

DAVEY, M. J., ANDRIGHETTI, H. J., MA, X. & BRANDL, C. J. 2011. A synthetic human kinase can control cell cycle progression in budding yeast. *G3 (Bethesda),* 1**,** 317-25.

DAY, T. A., PALLE, K., BARKLEY, L. R., KAKUSHO, N., ZOU, Y., TATEISHI, S., VERREAULT, A., MASAI, H. & VAZIRI, C. 2010. Phosphorylated Rad18 directs DNA polymerase eta to sites of stalled replication. *J Cell Biol,* 191**,** 953-66.

DE BONDT, H. L., ROSENBLATT, J., JANCARIK, J., JONES, H. D., MORGAN, D. O. & KIM, S. H. 1993. Crystal structure of cyclin-dependent kinase 2. *Nature,* 363**,** 595-602.

DEEGAN, T. D., YEELES, J. T. & DIFFLEY, J. F. 2016. Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation. *EMBO J,* 35**,** 961-73.

DIEDRICHS, K. & KARPLUS, P. A. 1997. Improved R-factors for diffraction data analysis in macromolecular crystallography (vol 4, pg 269, 1997). *Nature Structural Biology,* 4**,** 592-592.

DIEROV, J., DIEROVA, R. & CARROLL, M. 2004. BCR/ABL translocates to the nucleus and disrupts an ATR-dependent intra-S phase checkpoint. *Cancer Cell,* 5**,** 275-85.

DIFFLEY, J. F., COCKER, J. H., DOWELL, S. J. & ROWLEY, A. 1994. Two steps in the assembly of complexes at yeast replication origins in vivo. *Cell,* 78**,** 303-16.

DIMITROVA, D. S. & GILBERT, D. M. 1999. The spatial position and replication timing of chromosomal domains are both established in early G1 phase. *Mol Cell,* 4**,** 983-93.

DOHRMANN, P. R., OSHIRO, G., TECKLENBURG, M. & SCLAFANI, R. A. 1999. RAD53 regulates DBF4 independently of checkpoint function in Saccharomyces cerevisiae. *Genetics,* 151**,** 965-977.

DONALDSON, A. D., FANGMAN, W. L. & BREWER, B. J. 1998. Cdc7 is required throughout the yeast S phase to activate replication origins. *Genes & Development,* 12**,** 491-501.

DRURY, L. S., PERKINS, G. & DIFFLEY, J. F. X. 2000. The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. *Current Biology,* 10**,** 231-240.

DUNCKER, B. P., SHIMADA, K., TSAI-PFLUGFELDER, M., PASERO, P. & GASSER, S. M. 2002. An N-terminal domain of Dbf4p mediates interaction with both origin recognition complex (ORC) and Rad53p and can deregulate late origin firing. *Proc Natl Acad Sci U S A,* 99**,** 16087-92.

EATON, M. L., GALANI, K., KANG, S., BELL, S. P. & MACALPINE, D. M. 2010. Conserved nucleosome positioning defines replication origins. *Genes & Development,* 24**,** 748-753.

EBERHARTER, A. & BECKER, P. B. 2002. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO Rep,* 3**,** 224-9.

ECHALIER, A., ENDICOTT, J. A. & NOBLE, M. E. 2010. Recent developments in cyclin-dependent kinase biochemical and structural studies. *Biochim Biophys Acta,* 1804**,** 511-9.

EMSLEY, P. & COWTAN, K. 2004a. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D,* 60**,** 2126-2132.

EMSLEY, P. & COWTAN, K. 2004b. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography,* 60**,** 2126-2132.

ENDICOTT, J. A., NOBLE, M. E. & JOHNSON, L. N. 2012. The structural basis for control of eukaryotic protein kinases. *Annu Rev Biochem,* 81**,** 587-613.

ESER, U., FALLEUR-FETTIG, M., JOHNSON, A. & SKOTHEIM, J. M. 2011. Commitment to a Cellular Transition Precedes Genome-wide Transcriptional Change. *Molecular Cell,* 43**,** 515-527.

EVANS, P. 2006a. Scaling and assessment of data quality. *Acta Crystallographica Section D-Biological Crystallography,* 62**,** 72-82.

EVANS, P. 2006b. Scaling and assessment of data quality. *Acta Crystallographica Section D,* 62**,** 72-82.

EVRIN, C., CLARKE, P., ZECH, J., LURZ, R., SUN, J. C., UHLE, S., LI, H. L., STILLMAN, B. & SPECK, C. 2009. A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 20240-20245.

FAUL, T., STAIB, C., NANDA, I., SEHMID, M. & GRUMMT, F. 1999. Identification and characterization of mouse homologue to yeast Cdc7 protein and chromosomal localization of the cognate mouse gene Cdc71. *Chromosoma,* 108**,** 26-31.

FELDMAN, R. M. R., CORRELL, C. C., KAPLAN, K. B. & DESHAIES, R. J. 1997. A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell,* 91**,** 221-230.

FERENBACH, A., LI, A., BRITO-MARTINS, M. & BLOW, J. J. 2005. Functional domains of the Xenopus replication licensing factor Cdt1. *Nucleic Acids Res,* 33**,** 316-24.

FERNANDEZ-CID, A., RIERA, A., TOGNETTI, S., HERRERA, M. C., SAMEL, S., EVRIN, C., WINKLER, C., GARDENAL, E., UHLE, S. & SPECK, C. 2013. An ORC/Cdc6/MCM2-7 Complex Is Formed in a Multistep Reaction to Serve as a Platform for MCM Double-Hexamer Assembly. *Molecular Cell,* 50**,** 577-588.

FERREIRA, M. G., SANTOCANALE, C., DRURY, L. S. & DIFFLEY, J. F. X. 2000. Dbf4p, an essential S phase-promoting factor, is targeted for degradation by the anaphase-promoting complex. *Molecular and Cellular Biology,* 20**,** 242-248.

FRANCIS, L. I., RANDELL, J. C. W., TAKARA, T. J., UCHIMA, L. & BELL, S. P. 2009. Incorporation into the prereplicative complex activates the Mcm2-7 helicase for Cdc7-Dbf4 phosphorylation. *Genes & Development,* 23**,** 643-654.

FRIGOLA, J., REMUS, D., MEHANNA, A. & DIFFLEY, J. F. X. 2013. ATPase-dependent quality control of DNA replication origin licensing. *Nature,* 495**,** 339-343.

FULLER, R. S. & KORNBERG, A. 1983. Purified Dnaa Protein in Initiation of Replication at the Escherichia-Coli Chromosomal Origin of Replication. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences,* 80**,** 5817-5821.

GAGGIOLI, V., ZEISER, E., RIVERS, D., BRADSHAW, C. R., AHRINGER, J. & ZEGERMAN, P. 2014. CDK phosphorylation of SLD-2 is required for replication initiation and germline development in C. elegans. *J Cell Biol,* 204**,** 507-22.

GAMBUS, A., JONES, R. C., SANCHEZ-DIAZ, A., KANEMAKI, M., VAN DEURSEN, F., EDMONDSON, R. D. & LABIB, K. 2006. GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat Cell Biol,* 8**,** 358-66.

GAMBUS, A., KHOUDOLI, G. A., JONES, R. C. & BLOW, J. J. 2011. MCM2-7 Form Double Hexamers at Licensed Origins in Xenopus Egg Extract. *Journal of Biological Chemistry,* 286**,** 11855-11864.

GAO, Y., YAO, J., POUDEL, S., ROMER, E., ABU-NIAAJ, L. & LEFFAK, M. 2014. Protein phosphatase 2A and Cdc7 kinase regulate the DNA unwinding element-binding protein in replication initiation. *J Biol Chem,* 289**,** 35987-6000.

GERLITS, O., WALTMAN, M. J., TAYLOR, S., LANGAN, P. & KOVALEVSKY, A. 2013. Insights into the phosphoryl transfer catalyzed by cAMP-dependent protein kinase: an X-ray crystallographic study of complexes with various metals and peptide substrate SP20. *Biochemistry,* 52**,** 3721-7.

GHATALIA, P., YANG, E. S., LASSEIGNE, B. N., RAMAKER, R. C., COOPER, S. J., CHEN, D., SUDARSHAN, S., WEI, S., GURU, A. S., ZHAO, A., COOPER, T., DELLA MANNA, D. L., NAIK, G., MYERS, R. M. & SONPAVDE, G. 2016. Kinase Gene Expression Profiling of Metastatic Clear Cell Renal Cell

Carcinoma Tissue Identifies Potential New Therapeutic Targets. *PLoS One,* 11**,** e0160924.

GILLESPIE, P. J. & HIRANO, T. 2004. Scc2 couples replication licensing to sister chromatid cohesion in Xenopus egg extracts. *Curr Biol,* 14**,** 1598-603.

GONZALEZ, F. A., RADEN, D. L. & DAVIS, R. J. 1991. Identification of substrate recognition determinants for human ERK1 and ERK2 protein kinases. *J Biol Chem,* 266**,** 22159-63.

GOPALAKRISHNAN, V., SIMANCEK, P., HOUCHENS, C., SNAITH, H. A., FRATTINI, M. G., SAZER, S. & KELLY, T. J. 2001. Redundant control of rereplication in fission yeast. *Proc Natl Acad Sci U S A,* 98**,** 13114-9.

GUO, B. Q. & LEE, H. 1999. Cloning and characterization of Chinese hamster CDC7 (ChCDC7). *Somatic Cell and Molecular Genetics,* 25**,** 159-171.

HAAS, D. J. & ROSSMANN, M. G. 1970. Crystallographic Studies on Lactate Dehydrogenase at-75 Degrees C. *Acta Crystallographica Section B-Structural Crystallography and Crystal Chemistry,* B 26**,** 998-&.

HANKS, S. K., QUINN, A. M. & HUNTER, T. 1988. The Protein-Kinase Family - Conserved Features and Deduced Phylogeny of the Catalytic Domains. *Science,* 241**,** 42-52.

HARDY, C. F. J., DRYGA, O., SEEMATTER, S., PAHL, P. M. B. & SCLAFANI, R. A. 1997. mcm5/cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p. *Proceedings of the National Academy of Sciences of the United States of America,* 94**,** 3151-3155.

HARKINS, V., GABRIELSE, C., HASTE, L. & WEINREICH, M. 2009. Budding yeast Dbf4 sequences required for Cdc7 kinase activation and identification of a functional relationship between the Dbf4 and Rev1 BRCT domains. *Genetics,* 183**,** 1269-82.

HARTWELL, L. H. 1971. Genetic Control of Cell Division Cycle in Yeast .2. Genes Controlling DNA Replication and Its Initiation. *Journal of Molecular Biology,* 59**,** 183-&.

HARTWELL, L. H. 1973. 3 Additional Genes Required for Deoxyribonucleic Acid Synthesis in Saccharomyces-Cerevisiae. *Journal of Bacteriology,* 115**,** 966-974.

HAWKINS, J., ZHENG, S., FRANTZ, B. & LOGRASSO, P. 2000. p38 map kinase substrate specificity differs greatly for protein and peptide substrates. *Arch Biochem Biophys,* 382**,** 310-3.

HELLER, R. C., KANG, S., LAM, W. M., CHEN, S., CHAN, C. S. & BELL, S. P. 2011. Eukaryotic Origin-Dependent DNA Replication In Vitro Reveals Sequential Action of DDK and S-CDK Kinases. *Cell,* 146**,** 80-91.

HENDERSON, R. 1990. Cryoprotection of Protein Crystals against Radiation-Damage in Electron and X-Ray-Diffraction. *Proceedings of the Royal Society B-Biological Sciences,* 241**,** 6-8.

HEREFORD, L. M. & HARTWELL, L. H. 1974. Sequential Gene Function in Initiation of Saccharomyces-Cerevisiae DNA-Synthesis. *Journal of Molecular Biology,* 84**,** 445-&.

HINES, A. C., PARANG, K., KOHANSKI, R. A., HUBBARD, S. R. & COLE, P. A. 2005. Bisubstrate analog probes for the insulin receptor protein tyrosine kinase: Molecular yardsticks for analyzing catalytic mechanism and inhibitor design. *Bioorganic Chemistry,* 33**,** 285-297.

HIRAGA, S., ALVINO, G. M., CHANG, F., LIAN, H. Y., SRIDHAR, A., KUBOTA, T., BREWER, B. J., WEINREICH, M., RAGHURAMAN, M. K. & DONALDSON, A. D. 2014. Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex. *Genes Dev,* 28**,** 372-83.

HOLLINGSWORTH, R. E. & SCLAFANI, R. A. 1990. DNA Metabolism Gene Cdc7 from Yeast Encodes a Serine (Threonine) Protein-Kinase. *Proceedings of the National Academy of Sciences of the United States of America,* 87**,** 6272-6276.

HOSING, A. S., VALERIE, N. C., DZIEGIELEWSKI, J., BRAUTIGAN, D. L. & LARNER, J. M. 2012. PP6 regulatory subunit R1 is bidentate anchor for targeting protein phosphatase-6 to DNA-dependent protein kinase. *J Biol Chem,* 287**,** 9230-9.

HOU, Y., WANG, H. Q. & BA, Y. 2012. High expression of cell division cycle 7 protein correlates with poor prognosis in patients with diffuse large B-cell lymphoma. *Med Oncol,* 29**,** 3498-503.

HU, J. & XIONG, Y. 2006. An evolutionarily conserved function of proliferating cell nuclear antigen for Cdt1 degradation by the Cul4-Ddb1 ubiquitin ligase in response to DNA damage. *J Biol Chem,* 281**,** 3753-6.

HUANG, J., HUEN, M. S., KIM, H., LEUNG, C. C., GLOVER, J. N., YU, X. & CHEN, J. 2009. RAD18 transmits DNA damage signalling to elicit homologous recombination repair. *Nat Cell Biol,* 11**,** 592-603.

HUBBARD, S. R., BISHOP, W. R., KIRSCHMEIER, P., GEORGE, S. J., CRAMER, S. P. & HENDRICKSON, W. A. 1991. Identification and characterization of zinc binding sites in protein kinase C. *Science,* 254**,** 1776-9.

HUGGETT, M. T., TUDZAROVA, S., PROCTOR, I., LODDO, M., KEANE, M. G., STOEBER, K., WILLIAMS, G. H. & PEREIRA, S. P. 2016. Cdc7 is a potent anti-cancer target in pancreatic cancer due to abrogation of the DNA origin activation checkpoint. *Oncotarget,* 7**,** 18495-507.

HUGHES, S., ELUSTONDO, F., DI FONZO, A., LEROUX, F. G., WONG, A. C., SNIJDERS, A. P., MATTHEWS, S. J. & CHEREPANOV, P. 2012. Crystal structure of human CDC7 kinase in complex with its activator DBF4. *Nat Struct Mol Biol,* 19**,** 1101-7.

HUGHES, S., JENKINS, V., DAR, M. J., ENGELMAN, A. & CHEREPANOV, P. 2010. Transcriptional Co-activator LEDGF Interacts with Cdc7-Activator of S-phase Kinase (ASK) and Stimulates Its Enzymatic Activity. *Journal of Biological Chemistry,* 285**,** 541-554.

HUSE, M. & KURIYAN, J. 2002. The conformational plasticity of protein kinases. *Cell,* 109**,** 275-282.

ILVES, I., PETOJEVIC, T., PESAVENTO, J. J. & BOTCHAN, M. R. 2010. Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol Cell,* 37**,** 247-58.

IM, J. S., KI, S. H., FARINA, A., JUNG, D. S., HURWITZ, J. & LEE, J. K. 2009. Assembly of the Cdc45-Mcm2-7-GINS complex in human cells requires the Ctf4/And-1, RecQL4, and Mcm10 proteins. *Proc Natl Acad Sci U S A,* 106**,** 15628-32.

IM, J. S. & LEE, J. K. 2008. ATR-dependent activation of p38 MAP kinase is responsible for apoptotic cell death in cells depleted of Cdc7. *J Biol Chem,* 283**,** 25171-7.

IM, J. S., PARK, S. Y., CHO, W. H., BAE, S. H., HURWITZ, J. & LEE, J. K. 2015. RecQL4 is required for the association of Mcm10 and Ctf4 with replication origins in human cells. *Cell Cycle,* 14**,** 1001-9.

IRETON, G. C. & STODDARD, B. L. 2004. Microseed matrix screening to improve crystals of yeast cytosine deaminase. *Acta Crystallogr D Biol Crystallogr,* 60**,** 601-5.

ITO, S., ISHII, A., KAKUSHO, N., TANIYAMA, C., YAMAZAKI, S., FUKATSU, R., SAKAUE-SAWANO, A., MIYAWAKI, A. & MASAI, H. 2012. Mechanism of cancer cell death induced by depletion of an essential replication regulator. *PLoS One,* 7**,** e36372.

JACKSON, A. L., PAHL, P. M. B., HARRISON, K., ROSAMOND, J. & SCLAFANI, R. A. 1993. Cell-Cycle Regulation of the Yeast Cdc7 Protein-Kinase by Association with the Dbf4 Protein. *Molecular and Cellular Biology,* 13**,** 2899-2908.

JACOBSEN, D. M., BAO, Z. Q., O'BRIEN, P., BROOKS, C. L., 3RD & YOUNG, M. A. 2012. Price to be paid for two-metal catalysis: magnesium ions that accelerate chemistry unavoidably limit product release from a protein kinase. *J Am Chem Soc,* 134**,** 15357-70.

JARES, P. & BLOW, J. J. 2000. Xenopus Cdc7 function is dependent on licensing but not on XORC, XCdc6, or CDK activity and is required for XCdc45 loading. *Genes & Development,* 14**,** 1528-1540.

JARES, P., LUCIANI, M. G. & BLOW, J. J. 2004. A Xenopus Dbf4 homolog is required for Cdc7 chromatin binding and DNA replication. *BMC Mol Biol,* 5**,** 5.

JASPERSEN, S. L., CHARLES, J. F. & MORGAN, D. O. 1999. Inhibitory phosphorylation of the APC regulator Hct1 is controlled by the kinase Cdc28 and the phosphatase Cdc14. *Current Biology,* 9**,** 227-236.

JEFFREY, P. D., RUSO, A. A., POLYAK, K., GIBBS, E., HURWITZ, J., MASSAGUE, J. & PAVLETICH, N. P. 1995. Mechanism of Cdk Activation Revealed by the Structure of a Cyclina-Cdk2 Complex. *Nature,* 376**,** 313-320.

JIANG, W. & HUNTER, T. 1997. Identification and characterization of a human protein kinase related to budding yeast Cdc7p. *Proceedings of the National Academy of Sciences of the United States of America,* 94**,** 14320-14325.

JIANG, W., MCDONALD, D., HOPE, T. J. & HUNTER, T. 1999a. Mammalian Cdc7-Dbf4 protein kinase complex is essential for initiation of DNA replication. *Embo Journal,* 18**,** 5703-5713.

JIANG, W., WELLS, N. J. & HUNTER, T. 1999b. Multistep regulation of DNA replication by Cdk phosphorylation of HsCdc6. *Proc Natl Acad Sci U S A,* 96**,** 6193-8.

JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J. A. & CHARPENTIER, E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science,* 337**,** 816-21.

JOHNSTON, L. H. & THOMAS, A. P. 1982a. A Further 2 Mutants Defective in Initiation of the S-Phase in the Yeast Saccharomyces-Cerevisiae. *Molecular & General Genetics,* 186**,** 445-448.

JOHNSTON, L. H. & THOMAS, A. P. 1982b. The Isolation of New DNA-Synthesis Mutants in the Yeast Saccharomyces-Cerevisiae. *Molecular & General Genetics,* 186**,** 439-444.

KABSCH, W. 2010a. Xds. *Acta Crystallographica Section D-Biological Crystallography,* 66**,** 125-132.

KABSCH, W. 2010b. XDS. *Acta Crystallographica Section D,* 66**,** 125-132.

KALLUNKI, T., SU, B., TSIGELNY, I., SLUSS, H. K., DERIJARD, B., MOORE, G., DAVIS, R. & KARIN, M. 1994. JNK2 contains a specificity-determining region responsible for efficient c-Jun binding and phosphorylation. *Genes Dev,* 8**,** 2996-3007.

KANG, S., WARNER, M. D. & BELL, S. P. 2014. Multiple Functions for Mcm2-7 ATPase Motifs during Replication Initiation. *Molecular Cell,* 55**,** 655-665.

KANNAN, N., HASTE, N., TAYLOR, S. S. & NEUWALD, A. F. 2008. The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module (vol 104, pg 1272, 2007). *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 9130-9130.

KATIS, V. L., LIPP, J. J., IMRE, R., BOGDANOVA, A., OKAZ, E., HABERMANN, B., MECHTLER, K., NASMYTH, K. & ZACHARIAE, W. 2010. Rec8 phosphorylation by casein kinase 1 and Cdc7-Dbf4 kinase regulates cohesin cleavage by separase during meiosis. *Dev Cell,* 18**,** 397-409.

KAUFMANN, W. K., NEVIS, K. R., QU, P., IBRAHIM, J. G., ZHOU, T., ZHOU, Y., SIMPSON, D. A., HELMS-DEATON, J., CORDEIRO-STONE, M., MOORE, D. T., THOMAS, N. E., HAO, H., LIU, Z., SHIELDS, J. M., SCOTT, G. A. & SHARPLESS, N. E. 2008. Defective cell cycle checkpoint functions in melanoma are associated with altered patterns of gene expression. *J Invest Dermatol,* 128**,** 175-87.

KIHARA, M., NAKAI, W., ASANO, S., SUZUKI, A., KITADA, K., KAWASAKI, Y., JOHNSTON, L. H. & SUGINO, A. 2000. Characterization of the yeast Cdc7p/Dbf4p complex purified from insect cells. Its protein kinase activity is regulated by Rad53p. *J Biol Chem,* 275**,** 35051-62.

KIM, B. J., KIM, S. Y. & LEE, H. 2007. Identification and characterization of human Cdc7 nuclear retention and export sequences in the context of chromatin binding. *Journal of Biological Chemistry,* 282**,** 30029-30038.

KIM, B. J. & LEE, H. 2006. Importin-beta mediates Cdc7 nuclear import by binding to the kinase insert II domain, which can be antagonized by importin-alpha. *Journal of Biological Chemistry,* 281**,** 12041-12049.

KIM, J. M., KAKUSHO, N., YAMADA, M., KANOH, Y., TAKEMOTO, N. & MASAI, H. 2008. Cdc7 kinase mediates Claspin phosphorylation in DNA replication checkpoint. *Oncogene,* 27**,** 3475-82.

KIM, J. M., NAKAO, K., NAKAMURA, K., SAITO, I., KATSUKI, M., ARAI, K. & MASAI, H. 2002. Inactivation of Cdc7 kinase in mouse ES cells results in S-phase arrest and p53-dependent cell death. *Embo Journal,* 21**,** 2168-2179.

KIM, J. M., SATO, N., YAMADA, M., ARAI, K. & MASAI, H. 1998. Growth regulation of the expression of mouse cDNA and gene encoding a serine threonine kinase related to Saccharomyces cerevisiae CDC7 essential for G(1)/S transition. Structure, chromosomal localization, and expression of mouse gene for S-cerevisiae CDC7-related kinase. (vol 273, pg 23248, 1998). *Journal of Biological Chemistry,* 273**,** 27755-27755.

KITADA, K., JOHNSTON, L. H., SUGINO, T. & SUGINO, A. 1992. Temperature-Sensitive Cdc7 Mutations of Saccharomyces-Cerevisiae Are Suppressed by the Dbf4 Gene, Which Is Required for the G1/S Cell-Cycle Transition. *Genetics,* 131**,** 21-29.

KITAMURA, R., FUKATSU, R., KAKUSHO, N., CHO, Y. S., TANIYAMA, C., YAMAZAKI, S., TOH, G., YANAGI, K., ARAI, N., CHANG, H. J. & MASAI, H. 2011. Molecular Mechanism of Activation of Human Cdc7 Kinase BIPARTITE INTERACTION WITH Dbf4/ACTIVATOR OF S PHASE KINASE (ASK) ACTIVATION SUBUNIT STIMULATES ATP BINDING AND SUBSTRATE RECOGNITION. *Journal of Biological Chemistry,* 286**,** 23031-23043.

KLEMM, R. D., AUSTIN, R. J. & BELL, S. P. 1997. Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. *Cell,* 88**,** 493-502.

KNAPE, M. J., AHUJA, L. G., BERTINETTI, D., BURGHARDT, N. C., ZIMMERMANN, B., TAYLOR, S. S. & HERBERG, F. W. 2015. Divalent Metal Ions Mg(2)(+) and Ca(2)(+) Have Distinct Effects on Protein Kinase A Activity and Regulation. *ACS Chem Biol,* 10**,** 2303-15.

KNIGHTON, D. R., ZHENG, J. H., TEN EYCK, L. F., ASHFORD, V. A., XUONG, N. H., TAYLOR, S. S. & SOWADSKI, J. M. 1991a. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science,* 253**,** 407-14.

KNIGHTON, D. R., ZHENG, J. H., TEN EYCK, L. F., XUONG, N. H., TAYLOR, S. S. & SOWADSKI, J. M. 1991b. Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science,* 253**,** 414-20.

KNOCKLEBY, J., KIM, B. J., MEHTA, A. & LEE, H. 2016. Cdk1-mediated phosphorylation of Cdc7 suppresses DNA re-replication. *Cell Cycle,* 15**,** 1494-505.

KOHZAKI, M., CHIOUREA, M., VERSINI, G., ADACHI, N., TAKEDA, S., GAGOS, S. & HALAZONETIS, T. D. 2012. The helicase domain and C-terminus of human RecQL4 facilitate replication elongation on DNA templates damaged by ionizing radiation. *Carcinogenesis,* 33**,** 1203-10.

KOIVOMAGI, M., VALK, E., VENTA, R., IOFIK, A., LEPIKU, M., BALOG, E. R. M., RUBIN, S. M., MORGAN, D. O. & LOOG, M. 2011. Cascades of multisite phosphorylation control Sic1 destruction at the onset of S phase. *Nature,* 480**,** 128-U301.

KOLEK, S. A., BRAUNING, B. & STEWART, P. D. 2016. A novel microseeding method for the crystallization of membrane proteins in lipidic cubic phase. *Acta Crystallogr F Struct Biol Commun,* 72**,** 307-12.

KOLTUN, E. S., TSUHAKO, A. L., BROWN, D. S., AAY, N., ARCALAS, A., CHAN, V., DU, H., ENGST, S., FERGUSON, K., FRANZINI, M., GALAN, A., HOLST, C. R., HUANG, P., KANE, B., KIM, M. H., LI, J., MARKBY, D., MOHAN, M., NOSON, K., PLONOWSKI, A., RICHARDS, S. J., ROBERTSON, S., SHAW, K., STOTT, G., STOUT, T. J., YOUNG, J., YU, P., ZAHARIA, C. A., ZHANG, W., ZHOU, P., NUSS, J. M., XU, W. & KEARNEY, P. C. 2012. Discovery of XL413, a potent and selective CDC7 inhibitor. *Bioorg Med Chem Lett,* 22**,** 3727-31.

KORNEV, A. P. & TAYLOR, S. S. 2010. Defining the conserved internal architecture of a protein kinase. *Biochim Biophys Acta,* 1804**,** 440-4.

KORNEV, A. P., TAYLOR, S. S. & TEN EYCK, L. F. 2008. A helix scaffold for the assembly of active protein kinases. *Proc Natl Acad Sci U S A,* 105**,** 14377-82.

KRISHNA, S. S., MAJUMDAR, I. & GRISHIN, N. V. 2003. Structural classification of zinc fingers: survey and summary. *Nucleic Acids Res,* 31**,** 532-50.

KUBOTA, Y., TAKASE, Y., KOMORI, Y., HASHIMOTO, Y., ARATA, T., KAMIMURA, Y., ARAKI, H. & TAKISAWA, H. 2003. A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication. *Genes Dev,* 17**,** 1141-52.

KULKARNI, A. A., KINGSBURY, S. R., TUDZAROVA, S., HONG, H. K., LODDO, M., RASHID, M., RODRIGUEZ-ACEBES, S., PREVOST, A. T., LEDERMANN, J. A., STOEBER, K. & WILLIAMS, G. H. 2009. Cdc7 kinase is a predictor of survival and a novel therapeutic target in epithelial ovarian carcinoma. *Clin Cancer Res,* 15**,** 2417-25.

KUMAGAI, A., SHEVCHENKO, A., SHEVCHENKO, A. & DUNPHY, W. G. 2010. Treslin Collaborates with TopBP1 in Triggering the Initiation of DNA Replication. *Cell,* 140**,** 349-U59.

KUMAGAI, A., SHEVCHENKO, A., SHEVCHENKO, A. & DUNPHY, W. G. 2011. Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication. *Journal of Cell Biology,* 193**,** 995-1007.

KUMAGAI, H., SATO, N., YAMADA, M., MAHONY, D., SEGHEZZI, W., LEES, E., ARAI, K. I. & MASAI, H. 1999. A novel growth- and cell cycle-regulated protein, ASK, activates human Cdc7-related kinase and is essential for G(1)/S transition in mammalian cells. *Molecular and Cellular Biology,* 19**,** 5083-5095.

LAI, S. & PELECH, S. 2016. Regulatory roles of conserved phosphorylation sites in the activation T-loop of the MAP kinase ERK1. *Mol Biol Cell,* 27**,** 1040-50.

LAM, K. S. 1998. Determination of peptide substrate motifs for protein kinases using a "one-bead one-compound" combinatorial library approach. *Methods Mol Biol,* 87**,** 83-6.

LEE, A. Y., CHIBA, T., TRUONG, L. N., CHENG, A. N., DO, J., CHO, M. J., CHEN, L. & WU, X. 2012. Dbf4 is direct downstream target of ataxia telangiectasia

mutated (ATM) and ataxia telangiectasia and Rad3-related (ATR) protein to regulate intra-S-phase checkpoint. *J Biol Chem,* 287**,** 2531-43.
LEI, M., KAWASAKI, Y., YOUNG, M. R., KIHARA, M., SUGINO, A. & TYE, B. K. 1997. Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis. *Genes Dev,* 11**,** 3365-74.
LENGRONNE, A., MCINTYRE, J., KATOU, Y., KANOH, Y., HOPFNER, K. P., SHIRAHIGE, K. & UHLMANN, F. 2006. Establishment of sister chromatid cohesion at the S. cerevisiae replication fork. *Mol Cell,* 23**,** 787-99.
LESLIE, A. G. W. & POWELL, H. R. 2007. Processing diffraction data with MOSFLM. *Evolving Methods for Macromolecular Crystallography,* 245**,** 41-51.
LI, N., ZHAI, Y., ZHANG, Y., LI, W., YANG, M., LEI, J., TYE, B. K. & GAO, N. 2015a. Structure of the eukaryotic MCM complex at 3.8 A. *Nature,* 524**,** 186-91.
LI, W., ZHAO, X. L., SHANG, S. Q., SHEN, H. Q. & CHEN, X. 2015b. Dual Inhibition of Cdc7 and Cdk9 by PHA-767491 Suppresses Hepatocarcinoma Synergistically with 5-Fluorouracil. *Curr Cancer Drug Targets,* 15**,** 196-204.
LI, X. H., ZHAO, Q. P., LIAO, R., SUN, P. Q. & WU, X. H. 2003. The SCFSkp2 ubiquitin ligase complex interacts with the human replication licensing factor Cdt1 and regulates Cdt1 degradation. *Journal of Biological Chemistry,* 278**,** 30854-30858.
LINDBERG, R. A., QUINN, A. M. & HUNTER, T. 1992. Dual-specificity protein kinases: will any hydroxyl do? *Trends Biochem Sci,* 17**,** 114-9.
LO, H. C., KUNZ, R. C., CHEN, X., MARULLO, A., GYGI, S. P. & HOLLINGSWORTH, N. M. 2012. Cdc7-Dbf4 is a gene-specific regulator of meiotic transcription in yeast. *Mol Cell Biol,* 32**,** 541-57.
LO, H. C., WAN, L., ROSEBROCK, A., FUTCHER, B. & HOLLINGSWORTH, N. M. 2008. Cdc7-Dbf4 regulates NDT80 transcription as well as reductional segregation during budding yeast meiosis. *Mol Biol Cell,* 19**,** 4956-67.
LUO, K., ZHOU, P. & LODISH, H. F. 1995. The specificity of the transforming growth factor beta receptor kinases determined by a spatially addressable peptide library. *Proc Natl Acad Sci U S A,* 92**,** 11761-5.
MADHUSUDAN, TRAFNY, E. A., XUONG, N. H., ADAMS, J. A., TEN EYCK, L. F., TAYLOR, S. S. & SOWADSKI, J. M. 1994. cAMP-dependent protein kinase: crystallographic insights into substrate recognition and phosphotransfer. *Protein Sci,* 3**,** 176-87.
MAILAND, N. & DIFFLEY, J. F. 2005. CDKs promote DNA replication origin licensing in human cells by protecting Cdc6 from APC/C-dependent proteolysis. *Cell,* 122**,** 915-26.
MAKINIEMI, M., HILLUKKALA, T., TUUSA, J., REINI, K., VAARA, M., HUANG, D., POSPIECH, H., MAJURI, I., WESTERLING, T., MAKELA, T. P. & SYVAOJA, J. E. 2001. BRCT domain-containing protein TopBP1 functions in DNA replication and damage response. *J Biol Chem,* 276**,** 30399-406.
MALUMBRES, M. 2014. Cyclin-dependent kinases. *Genome Biol,* 15**,** 122.
MANNING, G., WHYTE, D. B., MARTINEZ, R., HUNTER, T. & SUDARSANAM, S. 2002. The protein kinase complement of the human genome. *Science,* 298**,** 1912-34.
MASAI, H. & ARAI, K. 2000. Dbf4 motifs: Conserved motifs in activation subunits for Cdc7 kinases essential for S-phase. *Biochemical and Biophysical Research Communications,* 275**,** 228-232.
MASAI, H., MATSUI, E., YOU, Z., ISHIMI, Y., TAMAI, K. & ARAI, K. 2000. Human Cdc7-related kinase complex. In vitro phosphorylation of MCM by concerted actions of Cdks and Cdc7 and that of a criticial threonine residue of Cdc7 bY Cdks. *J Biol Chem,* 275**,** 29042-52.

MASAI, H., MIYAKE, T. & ARAI, K. 1995. Hsk1(+), a Schizosaccharomyces-Pombe Gene-Related to Saccharomyces-Cerevisiae Cdc7, Is Required for Chromosomal Replication. *Embo Journal,* 14**,** 3094-3104.

MASAI, H., TANIYAMA, C., OGINO, K., MATSUI, E., KAKUSHO, N., MATSUMOTO, S., KIM, J. M., ISHII, A., TANAKA, T., KOBAYASHI, T., TAMAI, K., OHTANI, K. & ARAI, K. 2006. Phosphorylation of MCM4 by Cdc7 kinase facilitates its interaction with Cdc45 on the chromatin. *J Biol Chem,* 281**,** 39249-61.

MATOS, J., LIPP, J. J., BOGDANOVA, A., GUILLOT, S., OKAZ, E., JUNQUEIRA, M., SHEVCHENKO, A. & ZACHARIAE, W. G. 2008. Dbf4-Dependent Cdc7 Kinase Links DNA Replication to the Segregation of Homologous Chromosomes in Meiosis I. *Cell,* 135**,** 662-678.

MATSUNO, K., KUMANO, M., KUBOTA, Y., HASHIMOTO, Y. & TAKISAWA, H. 2006. The N-terminal noncatalytic region of Xenopus RecQ4 is required for chromatin binding of DNA polymerase alpha in the initiation of DNA replication. *Mol Cell Biol,* 26**,** 4843-52.

MATTAROCCI, S., SHYIAN, M., LEMMENS, L., DAMAY, P., ALTINTAS, D. M., SHI, T., BARTHOLOMEW, C. R., THOMA, N. H., HARDY, C. F. & SHORE, D. 2014. Rif1 controls DNA replication timing in yeast through the PP1 phosphatase Glc7. *Cell Rep,* 7**,** 62-9.

MATTHEWS, B. W. 1968. Solvent Content of Protein Crystals. *Journal of Molecular Biology,* 33**,** 491-&.

MATTHEWS, L. A., JONES, D. R., PRASAD, A. A., DUNCKER, B. P. & GUARNE, A. 2012. Saccharomyces cerevisiae Dbf4 has unique fold necessary for interaction with Rad53 kinase. *J Biol Chem,* 287**,** 2378-87.

MCCOY, A. J. 2007. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallographica Section D-Biological Crystallography,* 63**,** 32-41.

MCCOY, A. J., GROSSE-KUNSTLEVE, R. W., ADAMS, P. D., WINN, M. D., STORONI, L. C. & READ, R. J. 2007. Phaser crystallographic software. *Journal of Applied Crystallography,* 40**,** 658-674.

MCGARRY, T. J. & KIRSCHNER, M. W. 1998. Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell,* 93**,** 1043-1053.

MELLING, N., MUTH, J., SIMON, R., BOKEMEYER, C., TERRACCIANO, L., SAUTER, G., IZBICKI, J. R. & MARX, A. H. 2015. Cdc7 overexpression is an independent prognostic marker and a potential therapeutic target in colorectal cancer. *Diagn Pathol,* 10**,** 125.

MIHAYLOV, I. S., KONDO, T., JONES, L., RYZHIKOV, S., TANAKA, J., ZHENG, J., HIGA, L. A., MINAMINO, N., COOLEY, L. & ZHANG, H. 2002. Control of DNA replication and chromosome ploidy by geminin and cyclin A. *Mol Cell Biol,* 22**,** 1868-80.

MILLER, C. T., GABRIELSE, C., CHEN, Y. C. & WEINREICH, M. 2009. Cdc7p-Dbf4p Regulates Mitotic Exit by Inhibiting Polo Kinase. *Plos Genetics,* 5.

MIMURA, S., SEKI, T., TANAKA, S. & DIFFLEY, J. F. X. 2004. Phosphorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control. *Nature,* 431**,** 1118-1123.

MISHRA, P. K., CIFTCI-YILMAZ, S., REYNOLDS, D., AU, W. C., BOECKMANN, L., DITTMAN, L. E., JOWHAR, Z., PACHPOR, T., YEH, E., BAKER, R. E., HOYT, M. A., D'AMOURS, D., BLOOM, K. & BASRAI, M. A. 2016. Polo kinase Cdc5 associates with centromeres to facilitate the removal of centromeric cohesin during mitosis. *Mol Biol Cell,* 27**,** 2286-300.

MONTAGNOLI, A., BOSOTTI, R., VILLA, F., RIALLAND, M., BROTHERTON, D., MERCURIO, C., BERTHELSEN, J. & SANTOCANALE, C. 2002. Drf1, a novel regulatory subunit for human Cdc7 kinase. *Embo Journal,* 21**,** 3171-3181.

MONTAGNOLI, A., TENCA, P., SOLA, F., CARPANI, D., BROTHERTON, D., ALBANESE, C. & SANTOCANALE, C. 2004. Cdc7 inhibition reveals a p53-dependent replication checkpoint that is defective in cancer cells. *Cancer Research,* 64**,** 7110-7116.

MONTAGNOLI, A., VALSASINA, B., BROTHERTON, D., TROIANI, S., RAINOLDI, S., TENCA, P., MOLINARI, A. & SANTOCANALE, C. 2006. Identification of Mcm2 phosphorylation sites by S-phase-regulating kinases. *Journal of Biological Chemistry,* 281**,** 10281-10290.

MONTAGNOLI, A., VALSASINA, B., CROCI, V., MENICHINCHERI, M., RAINOLDI, S., MARCHESI, V., TIBOLLA, M., TENCA, P., BROTHERTON, D., ALBANESE, C., PATTON, V., ALZANI, R., CIAVOLELLA, A., SOLA, F., MOLINARI, A., VOLPI, D., AVANZI, N., FIORENTINI, F., CATTONI, M., HEALY, S., BALLINARI, D., PESENTI, E., ISACCHI, A., MOLL, J., BENSIMON, A., VANOTTI, E. & SANTOCANALE, C. 2008. A Cdc7 kinase inhibitor restricts initiation of DNA replication and has antitumor activity. *Nature Chemical Biology,* 4**,** 357-365.

MORGAN, D. O. 1995. Principles of CDK regulation. *Nature,* 374**,** 131-4.

MORGENSTERN, J. P. & LAND, H. 1990. Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucleic Acids Res,* 18**,** 3587-96.

MOYER, S. E., LEWIS, P. W. & BOTCHAN, M. R. 2006. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc Natl Acad Sci U S A,* 103**,** 10236-41.

MURAKAMI, H. & KEENEY, S. 2014. Temporospatial coordination of meiotic DNA replication and recombination via DDK recruitment to replisomes. *Cell,* 158**,** 861-73.

MURAMATSU, S., HIRAI, K., TAK, Y. S., KAMIMURA, Y. & ARAKI, H. 2010. CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol epsilon, and GINS in budding yeast. *Genes & Development,* 24**,** 602-612.

MURSHUDOV, G. N., VAGIN, A. A. & DODSON, E. J. 1997. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D,* 53**,** 240-255.

NAGY, A. 2000. Cre recombinase: the universal reagent for genome tailoring. *Genesis,* 26**,** 99-109.

NALDINI, L., BLOMER, U., GALLAY, P., ORY, D., MULLIGAN, R., GAGE, F. H., VERMA, I. M. & TRONO, D. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science,* 272**,** 263-7.

NAMBIAR, S., MIRMOHAMMADSADEGH, A., HASSAN, M., MOTA, R., MARINI, A., ALAOUI, A., TANNAPFEL, A., HEGEMANN, J. H. & HENGGE, U. R. 2007. Identification and functional characterization of ASK/Dbf4, a novel cell survival gene in cutaneous melanoma with prognostic relevance. *Carcinogenesis,* 28**,** 2501-10.

NASMYTH, K. & DIRICK, L. 1991. The Role of Swi4 and Swi6 in the Activity of G1 Cyclins in Yeast. *Cell,* 66**,** 995-1013.

NATSUME, T., KIYOMITSU, T., SAGA, Y. & KANEMAKI, M. T. 2016. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep,* 15**,** 210-8.

NATSUME, T., MULLER, C. A., KATOU, Y., RETKUTE, R., GIERLINSKI, M., ARAKI, H., BLOW, J. J., SHIRAHIGE, K., NIEDUSZYNSKI, C. A. & TANAKA, T. U. 2013. Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. *Mol Cell,* 50**,** 661-74.

NGUYEN, V. Q., CO, C., IRIE, K. & LI, J. J. 2000. Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7. *Current Biology,* 10**,** 195-205.

NGUYEN, V. Q., CO, C. & LI, J. J. 2001. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature,* 411**,** 1068-1073.
NISHIKAWA, K., TOKER, A., JOHANNES, F. J., SONGYANG, Z. & CANTLEY, L. C. 1997. Determination of the specific substrate sequence motifs of protein kinase C isozymes. *J Biol Chem,* 272**,** 952-60.
NISHITANI, H., SUGIMOTO, N., ROUKOS, V., NAKANISHI, Y., SAIJO, M., OBUSE, C., TSURIMOTO, T., NAKAYAMA, K. I., NAKAYAMA, K., FUJITA, M., LYGEROU, Z. & NISHIMOTO, T. 2006. Two E3 ubiquitin ligases, SCF-Skp2 and DDB1-Cul4, target human Cdt1 for proteolysis. *Embo Journal,* 25**,** 1126-1136.
NJAGI, G. D. & KILBEY, B. J. 1982. cdc7-1 a temperature sensitive cell-cycle mutant which interferes with induced mutagenesis in Saccharomyces cerevisiae. *Mol Gen Genet,* 186**,** 478-81.
NOLEN, B., TAYLOR, S. & GHOSH, G. 2004. Regulation of protein kinases; controlling activity through activation segment conformation. *Mol Cell,* 15**,** 661-75.
NONAKA, N., KITAJIMA, T., YOKOBAYASHI, S., XIAO, G., YAMAMOTO, M., GREWAL, S. I. & WATANABE, Y. 2002. Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nat Cell Biol,* 4**,** 89-93.
NOUGAREDE, R., DELLA SETA, F., ZARZOV, P. & SCHWOB, E. 2000. Hierarchy of S-phase-promoting factors: yeast Dbf4-Cdc7 kinase requires prior S-phase cyclin-dependent kinase activation. *Mol Cell Biol,* 20**,** 3795-806.
OBMOLOVA, G., MALIA, T. J., TEPLYAKOV, A., SWEET, R. & GILLILAND, G. L. 2010. Promoting crystallization of antibody-antigen complexes via microseed matrix screening. *Acta Crystallogr D Biol Crystallogr,* 66**,** 927-33.
OCASIO, C. A., RAJASEKARAN, M. B., WALKER, S., LE GRAND, D., SPENCER, J., PEARL, F. M., WARD, S. E., SAVIC, V., PEARL, L. H., HOCHEGGER, H. & OLIVER, A. W. 2016. A first generation inhibitor of human Greatwall kinase, enabled by structural and functional characterisation of a minimal kinase domain construct. *Oncotarget*.
OGINO, K., HIROTA, K., MATSUMOTO, S., TAKEDA, T., OHTA, K., ARAI, K. & MASAI, H. 2006. Hsk1 kinase is required for induction of meiotic dsDNA breaks without involving checkpoint kinases in fission yeast. *Proc Natl Acad Sci U S A,* 103**,** 8131-6.
OGINO, K., TAKEDA, T., MATSUI, E., IIYAMA, H., TANIYAMA, C., ARAI, K. & MASAI, H. 2001. Bipartite binding of a kinase activator activates Cdc7-related kinase essential for S phase. *Journal of Biological Chemistry,* 276**,** 31376-31387.
ON, K. F., BEURON, F., FRITH, D., SNIJDERS, A. P., MORRIS, E. P. & DIFFLEY, J. F. X. 2014. Prereplicative complexes assembled in vitro support origin-dependent and independent DNA replication. *Embo Journal,* 33**,** 605-620.
OSHIRO, G., OWENS, J. C., SHELLMAN, Y., SCLAFANI, R. A. & LI, J. J. 1999. Cell cycle control of Cdc7p kinase activity through regulation of Dbf4p stability. *Molecular and Cellular Biology,* 19**,** 4888-4896.
PACE, N. J. & WEERAPANA, E. 2014. Zinc-binding cysteines: diverse functions and structural motifs. *Biomolecules,* 4**,** 419-34.
PACEK, M. & WALTER, J. C. 2004. A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. *EMBO J,* 23**,** 3667-76.
PARANG, K., TILL, J. H., ABLOOGLU, A. J., KOHANSKI, R. A., HUBBARD, S. R. & COLE, P. A. 2001. Mechanism-based design of a protein kinase inhibitor. *Nat Struct Biol,* 8**,** 37-41.
PASERO, P., DUNCKER, B. P., SCHWOB, E. & GASSER, S. M. 1999. A role for the Cdc7 kinase regulatory subunit Dbf4p in the formation of initiation-competent origins of replication. *Genes Dev,* 13**,** 2159-76.

PATEL, P. K., KOMMAJOSYULA, N., ROSEBROCK, A., BENSIMON, A., LEATHERWOOD, J., BECHHOEFER, J. & RHIND, N. 2008. The Hsk1(Cdc7) replication kinase regulates origin efficiency. *Mol Biol Cell,* 19**,** 5550-8.

PATTERSON, M., SCLAFANI, R. A., FANGMAN, W. L. & ROSAMOND, J. 1986. Molecular Characterization of Cell-Cycle Gene Cdc7 from Saccharomyces-Cerevisiae. *Molecular and Cellular Biology,* 6**,** 1590-1598.

PEARSON, R. B. & KEMP, B. E. 1991. Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulations. *Methods Enzymol,* 200**,** 62-81.

PESSOA-BRANDAO, L. & SCLAFANI, R. A. 2004. CDC7/DBF4 functions in the translesion synthesis branch of the RAD6 epistasis group in Saccharomyces cerevisiae. *Genetics,* 167**,** 1597-610.

POH, W. T., CHADHA, G. S., GILLESPIE, P. J., KALDIS, P. & BLOW, J. J. 2014. Xenopus Cdc7 executes its essential function early in S phase and is counteracted by checkpoint-regulated protein phosphatase 1. *Open Biol,* 4**,** 130138.

PTACEK, J., DEVGAN, G., MICHAUD, G., ZHU, H., ZHU, X., FASOLO, J., GUO, H., JONA, G., BREITKREUTZ, A., SOPKO, R., MCCARTNEY, R. R., SCHMIDT, M. C., RACHIDI, N., LEE, S. J., MAH, A. S., MENG, L., STARK, M. J., STERN, D. F., DE VIRGILIO, C., TYERS, M., ANDREWS, B., GERSTEIN, M., SCHWEITZER, B., PREDKI, P. F. & SNYDER, M. 2005. Global analysis of protein phosphorylation in yeast. *Nature,* 438**,** 679-84.

RAN, F. A., HSU, P. D., WRIGHT, J., AGARWALA, V., SCOTT, D. A. & ZHANG, F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc,* 8**,** 2281-308.

RANDELL, J. C. W., FAN, A., CHAN, C., FRANCIS, L. I., HELLER, R. C., GALANI, K. & BELL, S. P. 2010. Mec1 Is One of Multiple Kinases that Prime the Mcm2-7 Helicase for Phosphorylation by Cdc7. *Molecular Cell,* 40**,** 353-363.

RAO, P. N. & JOHNSON, R. T. 1970. Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis. *Nature,* 225**,** 159-64.

RASHEED, S., NELSON-REES, W. A., TOTH, E. M., ARNSTEIN, P. & GARDNER, M. B. 1974. Characterization of a newly derived human sarcoma cell line (HT-1080). *Cancer,* 33**,** 1027-33.

RELLOS, P., IVINS, F. J., BAXTER, J. E., PIKE, A., NOTT, T. J., PARKINSON, D. M., DAS, S., HOWELL, S., FEDOROV, O., SHEN, Q. Y., FRY, A. M., KNAPP, S. & SMERDON, S. J. 2007. Structure and regulation of the human Nek2 centrosomal kinase. *J Biol Chem,* 282**,** 6833-42.

REMUS, D., BEALL, E. L. & BOTCHAN, M. R. 2004. DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. *Embo Journal,* 23**,** 897-907.

REMUS, D., BEURON, F., TOLUN, G., GRIFFITH, J. D., MORRIS, E. P. & DIFFLEY, J. F. 2009. Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell,* 139**,** 719-30.

REMUS, D. & DIFFLEY, J. F. 2009. Eukaryotic DNA replication control: lock and load, then fire. *Curr Opin Cell Biol,* 21**,** 771-7.

RICOUART, A., GESQUIERE, J. C., TARTAR, A. & SERGHERAERT, C. 1991. Design of potent protein kinase inhibitors using the bisubstrate approach. *J Med Chem,* 34**,** 73-8.

RUPP, B. & KANTARDJIEFF, K. 2010. *Biomolecular crystallography : principles, practice, and application to structural biology,* New York, Garland Science.

RUSSO, A. A., JEFFREY, P. D. & PAVLETICH, N. P. 1996. Structural basis of cyclin-dependent kinase activation by phosphorylation. *Nature Structural Biology,* 3**,** 696-700.

SAKUNO, T. & WATANABE, Y. 2009. Studies of meiosis disclose distinct roles of cohesion in the core centromere and pericentromeric regions. *Chromosome Res,* 17**,** 239-49.

SAMSON, R. Y., XU, Y., GADELHA, C., STONE, T. A., FAQIRI, J. N., LI, D., QIN, N., PU, F., LIANG, Y. X., SHE, Q. & BELL, S. D. 2013. Specificity and function of archaeal DNA replication initiator proteins. *Cell Rep,* 3**,** 485-96.

SANCHEZ-PULIDO, L., DIFFLEY, J. F. & PONTING, C. P. 2010. Homology explains the functional similarities of Treslin/Ticrr and Sld3. *Curr Biol,* 20**,** R509-10.

SANTOCANALE, C. & DIFFLEY, J. F. 1998. A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. *Nature,* 395**,** 615-8.

SASANUMA, H., HIROTA, K., FUKUDA, T., KAKUSHO, N., KUGOU, K., KAWASAKI, Y., SHIBATA, T., MASAI, H. & OHTA, K. 2008. Cdc7-dependent phosphorylation of Mer2 facilitates initiation of yeast meiotic recombination. *Genes & Development,* 22**,** 398-410.

SASI, N. K., TIWARI, K., SOON, F. F., BONTE, D., WANG, T., MELCHER, K., XU, H. E. & WEINREICH, M. 2014. The potent Cdc7-Dbf4 (DDK) kinase inhibitor XL413 has limited activity in many cancer cell lines and discovery of potential new DDK inhibitor scaffolds. *PLoS One,* 9**,** e113300.

SATO, N., ARAI, K. & MASAI, H. 1997. Human and Xenopus cDNA encoding budding yeast Cdc7-related kinases: in vitro phosphorylation of MCM subunits by a putative human homologue of Cdc7. *Embo Journal,* 16**,** 4340-4351.

SATO, N., SATO, M., NAKAYAMA, M., SAITOH, R., ARAI, K. & MASAI, H. 2003. Cell cycle regulation of chromatin binding and nuclear localization of human Cdc7-ASK kinase complex. *Genes Cells,* 8**,** 451-63.

SAWA, M. & MASAI, H. 2009. Drug design with Cdc7 kinase: a potential novel cancer therapy target. *Drug Des Devel Ther,* 2**,** 255-64.

SCHAEFFER, H. J. & WEBER, M. J. 1999. Mitogen-activated protein kinases: specific messages from ubiquitous messengers. *Mol Cell Biol,* 19**,** 2435-44.

SCHNEIDER, T. L., MATHEW, R. S., RICE, K. P., TAMAKI, K., WOOD, J. L. & SCHEPARTZ, A. 2005. Increasing the kinase specificity of k252a by protein surface recognition. *Org Lett,* 7**,** 1695-8.

SCHULMAN, B. A., LINDSTROM, D. L. & HARLOW, E. 1998. Substrate recruitment to cyclin-dependent kinase 2 by a multipurpose docking site on cyclin A. *Proc Natl Acad Sci U S A,* 95**,** 10453-8.

SCHWARZ, J. K., LOVLY, C. M. & PIWNICA-WORMS, H. 2003. Regulation of the Chk2 protein kinase by oligomerization-mediated cis- and trans-phosphorylation. *Mol Cancer Res,* 1**,** 598-609.

SEFTON, B. M. & SHENOLIKAR, S. 2001. Overview of protein phosphorylation. *Curr Protoc Protein Sci,* Chapter 13**,** Unit13 1.

SEGURADO, M. & TERCERO, J. A. 2009. The S-phase checkpoint: targeting the replication fork. *Biol Cell,* 101**,** 617-27.

SESSA, F., MAPELLI, M., CIFERRI, C., TARRICONE, C., ARECES, L. B., SCHNEIDER, T. R., STUKENBERG, R. T. & MUSACCHIO, A. 2005. Mechanism of Aurora B activation by INCENP and inhibition by Hesperadin. *Molecular Cell,* 18**,** 379-391.

SHARROCKS, A. D., YANG, S. H. & GALANIS, A. 2000. Docking domains and substrate-specificity determination for MAP kinases. *Trends Biochem Sci,* 25**,** 448-53.

SHEU, Y. J. & STILLMAN, B. 2006. Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. *Molecular Cell,* 24**,** 101-113.

SHEU, Y. J. & STILLMAN, B. 2010. The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature,* 463**,** 113-7.

SILVA, T., BRADLEY, R. H., GAO, Y. & COUE, M. 2006. Xenopus CDC7/DRF1 complex is required for the initiation of DNA replication. *J Biol Chem,* 281**,** 11569-76.

SNAITH, H. A., BROWN, G. W. & FORSBURG, S. L. 2000. Schizosaccharomyces pombe Hsk1p is a potential cds1p target required for genome integrity. *Mol Cell Biol,* 20**,** 7922-32.

SONGYANG, Z., BLECHNER, S., HOAGLAND, N., HOEKSTRA, M. F., PIWNICA-WORMS, H. & CANTLEY, L. C. 1994. Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Curr Biol,* 4**,** 973-82.

SONGYANG, Z., CARRAWAY, K. L., 3RD, ECK, M. J., HARRISON, S. C., FELDMAN, R. A., MOHAMMADI, M., SCHLESSINGER, J., HUBBARD, S. R., SMITH, D. P., ENG, C. & ET AL. 1995. Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature,* 373**,** 536-9.

SONGYANG, Z., LU, K. P., KWON, Y. T., TSAI, L. H., FILHOL, O., COCHET, C., BRICKEY, D. A., SODERLING, T. R., BARTLESON, C., GRAVES, D. J., DEMAGGIO, A. J., HOEKSTRA, M. F., BLENIS, J., HUNTER, T. & CANTLEY, L. C. 1996. A structural basis for substrate specificities of protein Ser/Thr kinases: primary sequence preference of casein kinases I and II, NIMA, phosphorylase kinase, calmodulin-dependent kinase II, CDK5, and Erk1. *Mol Cell Biol,* 16**,** 6486-93.

SPECK, C., CHEN, Z. Q., LI, H. L. & STILLMAN, B. 2005. ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nature Structural & Molecular Biology,* 12**,** 965-971.

SPELLMAN, P. T., SHERLOCK, G., ZHANG, M. Q., IYER, V. R., ANDERS, K., EISEN, M. B., BROWN, P. O., BOTSTEIN, D. & FUTCHER, B. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Molecular Biology of the Cell,* 9**,** 3273-3297.

STEBBINS, J. L., DE, S. K., PAVLICKOVA, P., CHEN, V., MACHLEIDT, T., CHEN, L. H., KUNTZEN, C., KITADA, S., KARIN, M. & PELLECCHIA, M. 2011. Design and characterization of a potent and selective dual ATP- and substrate-competitive subnanomolar bidentate c-Jun N-terminal kinase (JNK) inhibitor. *J Med Chem,* 54**,** 6206-14.

STEFANSSON, B. & BRAUTIGAN, D. L. 2007. Protein phosphatase PP6 N terminal domain restricts G1 to S phase progression in human cancer cells. *Cell Cycle,* 6**,** 1386-92.

STERNBERG, S. H., REDDING, S., JINEK, M., GREENE, E. C. & DOUDNA, J. A. 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature,* 507**,** 62-7.

STEVENSON-LINDERT, L. M., FOWLER, P. & LEW, J. 2003. Substrate specificity of CDK2-cyclin A. What is optimal? *J Biol Chem,* 278**,** 50956-60.

TAKAHASHI, T. S., BASU, A., BERMUDEZ, V., HURWITZ, J. & WALTER, J. C. 2008. Cdc7-Drf1 kinase links chromosome cohesion to the initiation of DNA replication in Xenopus egg extracts. *Genes & Development,* 22**,** 1894-1905.

TAKAHASHI, T. S. & WALTER, J. C. 2005a. Cdc7-Drf1 is a developmentally regulated protein kinase required for the initiation of vertebrate DNA replication. *Genes & Development,* 19**,** 2295-2300.

TAKAHASHI, T. S. & WALTER, J. C. 2005b. Cdc7-Drf1 is a developmentally regulated protein kinase required for the initiation of vertebrate DNA replication. *Genes Dev,* 19**,** 2295-300.

TAKAHASHI, T. S., YIU, P., CHOU, M. F., GYGI, S. & WALTER, J. C. 2004. Recruitment of Xenopus Scc2 and cohesin to chromatin requires the pre-replication complex. *Nat Cell Biol,* 6**,** 991-6.

TAKEDA, T., OGINO, K., MATSUI, E., CHO, M. K., KUMAGAI, H., MIYAKE, T., ARAI, K. & MASAI, H. 1999. A fission yeast gene, him1(+)/dfp1(+), encoding a regulatory subunit for Hsk1 kinase, plays essential roles in S-phase initiation as well as in S-phase checkpoint control and recovery from DNA damage. *Molecular and Cellular Biology,* 19**,** 5535-5547.

TAKEDA, T., OGINO, K., TATEBAYASHI, K., IKEDA, H., ARAI, K. & MASAI, H. 2001. Regulation of initiation of S phase, replication checkpoint signaling, and maintenance of mitotic chromosome structures during S phase by Hsk1 kinase in the fission yeast. *Mol Biol Cell,* 12**,** 1257-74.

TANAKA, S. & ARAKI, H. 2013. Helicase Activation and Establishment of Replication Forks at Chromosomal Origins of Replication. *Cold Spring Harbor Perspectives in Biology,* 5.

TANAKA, S. & DIFFLEY, J. F. X. 2002. Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. *Nature Cell Biology,* 4**,** 198-207.

TANAKA, S., UMEMORI, T., HIRAI, K., MURAMATSU, S., KAMIMURA, Y. & ARAKI, H. 2007. CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature,* 445**,** 328-332.

TANAKA, T., UMEMORI, T., ENDO, S., MURAMATSU, S., KANEMAKI, M., KAMIMURA, Y., OBUSE, C. & ARAKI, H. 2011. Sld7, an Sld3-associated protein required for efficient chromosomal DNA replication in budding yeast. *Embo Journal,* 30**,** 2019-2030.

TAYLOR, S. S. & KORNEV, A. P. 2011. Protein kinases: evolution of dynamic regulatory proteins. *Trends Biochem Sci,* 36**,** 65-77.

TENCA, P., BROTHERTON, D., MONTAGNOLI, A., RAINOLDI, S., ALBANESE, C. & SANTOCANALE, C. 2007. Cdc7 is an active kinase in human cancer cells undergoing replication stress. *J Biol Chem,* 282**,** 208-15.

TICAU, S., FRIEDMAN, L. J., IVICA, N. A., GELLES, J. & BELL, S. P. 2015. Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. *Cell,* 161**,** 513-25.

TILL, M., ROBSON, A., BYRNE, M. J., NAIR, A. V., KOLEK, S. A., SHAW STEWART, P. D. & RACE, P. R. 2013. Improving the success rate of protein crystallization by random microseed matrix screening. *J Vis Exp*.

TIMM, T., LI, X. Y., BIERNAT, J., JIAO, J., MANDELKOW, E., VANDEKERCKHOVE, J. & MANDELKOW, E. M. 2003. MARKK, a Ste20-like kinase, activates the polarity-inducing kinase MARK/PAR-1. *EMBO J,* 22**,** 5090-101.

TIMM, T., MARX, A., PANNEERSELVAM, S., MANDELKOW, E. & MANDELKOW, E. M. 2008. Structure and regulation of MARK, a kinase involved in abnormal phosphorylation of Tau protein. *BMC Neurosci,* 9 Suppl 2**,** S9.

TREIBER, D. K. & SHAH, N. P. 2013. Ins and outs of kinase DFG motifs. *Chem Biol,* 20**,** 745-6.

TSUJI, T., FICARRO, S. B. & JIANG, W. 2006. Essential role of phosphorylation of MCM2 by Cdc7/Dbf4 in the initiation of DNA replication in mammalian cells. *Mol Biol Cell,* 17**,** 4459-72.

TSUJI, T., LAU, E., CHIANG, G. G. & JIANG, W. 2008. The role of Dbf4/Drf1-dependent kinase Cdc7 in DNA-damage checkpoint control. *Mol Cell,* 32**,** 862-9.

TUDZAROVA, S., MULHOLLAND, P., DEY, A., STOEBER, K., OKOROKOV, A. L. & WILLIAMS, G. H. 2016. p53 controls CDC7 levels to reinforce G1 cell cycle arrest upon genotoxic stress. *Cell Cycle***,** 1-15.

UHLMANN, F. & NASMYTH, K. 1998. Cohesion between sister chromatids must be established during DNA replication. *Curr Biol,* 8**,** 1095-101.

ULM, J. W., PERRON, M., SODROSKI, J. & MULLIGAN, R. C. 2007. Complex determinants within the Moloney murine leukemia virus capsid modulate

susceptibility of the virus to Fv1 and Ref1-mediated restriction. *Virology,* 363**,** 245-255.

UNNIKRISHNAN, A., GAFKEN, P. R. & TSUKIYAMA, T. 2010. Dynamic changes in histone acetylation regulate origins of DNA replication. *Nat Struct Mol Biol,* 17**,** 430-7.

VALENTIN, G., SCHWOB, E. & DELLA SETA, F. 2006. Dual role of the Cdc7-regulatory protein Dbf4 during yeast meiosis. *J Biol Chem,* 281**,** 2828-34.

VAN AMEIJDE, J., POOT, A. J., VAN WANDELEN, L. T. M., WAMMES, A. E. M., RUIJTENBEEK, R., RIJKERS, D. T. S. & LISKAMP, R. M. J. 2010. Preparation of novel alkylated arginine derivatives suitable for click-cycloaddition chemistry and their incorporation into pseudosubstrate- and bisubstrate-based kinase inhibitors. *Organic & Biomolecular Chemistry,* 8**,** 1629-1639.

VAN DEURSEN, F., SENGUPTA, S., DE PICCOLI, G., SANCHEZ-DIAZ, A. & LABIB, K. 2012. Mcm10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation. *Embo Journal,* 31**,** 2195-2206.

VANOTTI, E., AMICI, R., BARGIOTTI, A., BERTHELSEN, J., BOSOTTI, R., CIAVOLELLA, A., CIRLA, A., CRISTIANI, C., D'ALESSIO, R., FORTE, B., ISACCHI, A., MARTINA, K., MENICHINCHERI, M., MOLINARI, A., MONTAGNOLI, A., ORSINI, P., PILLAN, A., ROLETTO, F., SCOLARO, A., TIBOLLA, M., VALSASINA, B., VARASI, M., VOLPI, D. & SANTOCANALE, C. 2008. Cdc7 kinase inhibitors: pyrrolopyridinones as potential antitumor agents. 1. Synthesis and structure-activity relationships. *J Med Chem,* 51**,** 487-501.

VARRIN, A. E., PRASAD, A. A., SCHOLZ, R. P., RAMER, M. D. & DUNCKER, B. P. 2005. A mutation in Dbf4 motif M impairs interactions with DNA replication factors and confers increased resistance to genotoxic agents. *Mol Cell Biol,* 25**,** 7494-504.

VAZIRI, C. & MASAI, H. 2010. Integrating DNA replication with trans-lesion synthesis via Cdc7. *Cell Cycle,* 9**,** 4818-23.

VERMA, R., ANNAN, R. S., HUDDLESTON, M. J., CARR, S. A., REYNARD, G. & DESHAIES, R. J. 1997. Phosphorylation of Sic1p by G(1) Cdk required for its degradation and entry into S phase. *Science,* 278**,** 455-460.

VIGNERON, S., BRIOUDES, E., BURGESS, A., LABBE, J. C., LORCA, T. & CASTRO, A. 2009. Greatwall maintains mitosis through regulation of PP2A. *EMBO J,* 28**,** 2786-93.

WALTER, J. C. 2000. Evidence for sequential action of cdc7 and cdk2 protein kinases during initiation of DNA replication in Xenopus egg extracts. *Journal of Biological Chemistry,* 275**,** 39773-39778.

WAN, L., NIU, H., FUTCHER, B., ZHANG, C., SHOKAT, K. M., BOULTON, S. J. & HOLLINGSWORTH, N. M. 2008. Cdc28-Clb5 (CDK-S) and Cdc7-Dbf4 (DDK) collaborate to initiate meiotic recombination in yeast. *Genes Dev,* 22**,** 386-97.

WEINREICH, M., LIANG, C., CHEN, H. H. & STILLMAN, B. 2001. Binding of cyclin-dependent kinases to ORC and Cdc6p regulates the chromosome replication cycle. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 11211-11217.

WEINREICH, M. & STILLMAN, B. 1999. Cdc7p-Dbf4p kinase binds to chromatin during S phase and is regulated by both the APC and the RAD53 checkpoint pathway. *Embo Journal,* 18**,** 5334-5346.

WEISS, M. S. & HILGENFELD, R. 1997. On the use of the merging R factor as a quality indicator for X-ray data. *Journal of Applied Crystallography,* 30**,** 203-205.

WILMES, G. M., ARCHAMBAULT, V., AUSTIN, R. J., JACOBSON, M. D., BELL, S. P. & CROSS, F. R. 2004. Interaction of the S-phase cyclin Clb5 with an 'RXL' docking sequence in the initiator protein Orc6 provides an origin-localized replication control switch. *Genes & Development,* 18**,** 981-991.

WINTER, G. 2010a. xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography,* 43**,** 186-190.
WINTER, G. 2010b. xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography,* 43**,** 186-190.
WOHLSCHLEGEL, J. A., DWYER, B. T., DHAR, S. K., CVETIC, C., WALTER, J. C. & DUTTA, A. 2000. Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science,* 290**,** 2309-+.
XU, X., ROCHETTE, P. J., FEYISSA, E. A., SU, T. V. & LIU, Y. 2009. MCM10 mediates RECQ4 association with MCM2-7 helicase complex during DNA replication. *EMBO J,* 28**,** 3005-14.
YABUUCHI, H., YAMADA, Y., UCHIDA, T., SUNATHVANICHKUL, T., NAKAGAWA, T. & MASUKATA, H. 2006. Ordered assembly of Sld3, GINS and Cdc45 is distinctly regulated by DDK and CDK for activation of replication origins. *Embo Journal,* 25**,** 4663-4674.
YAMADA, M., WATANABE, K., MISTRIK, M., VESELA, E., PROTIVANKOVA, I., MAILAND, N., LEE, M., MASAI, H., LUKAS, J. & BARTEK, J. 2013. ATR-Chk1-APC/C-Cdh1-dependent stabilization of Cdc7-ASK (Dbf4) kinase is required for DNA lesion bypass under replication stress. *Genes & Development,* 27**,** 2459-2472.
YAMAGUCHI, H. & HENDRICKSON, W. A. 1996. Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. *Nature,* 384**,** 484-9.
YAMAGUCHI, H., MATSUSHITA, M., NAIRN, A. C. & KURIYAN, J. 2001. Crystal structure of the atypical protein kinase domain of a TRP channel with phosphotransferase activity. *Mol Cell,* 7**,** 1047-57.
YAMAZAKI, S., ISHII, A., KANOH, Y., ODA, M., NISHITO, Y. & MASAI, H. 2012. Rif1 regulates the replication timing domains on the human genome. *EMBO J,* 31**,** 3667-77.
YANG, C. C., SUZUKI, M., YAMAKAWA, S., UNO, S., ISHII, A., YAMAZAKI, S., FUKATSU, R., FUJISAWA, R., SAKIMURA, K., TSURIMOTO, T. & MASAI, H. 2016. Claspin recruits Cdc7 kinase for initiation of DNA replication in human cells. *Nat Commun,* 7**,** 12135.
YANG, J., CRON, P., THOMPSON, V., GOOD, V. M., HESS, D., HEMMINGS, B. A. & BARFORD, D. 2002. Molecular mechanism for the regulation of protein kinase B/Akt by hydrophobic motif phosphorylation. *Molecular Cell,* 9**,** 1227-1240.
YARDIMCI, H. & WALTER, J. C. 2014. Prereplication-complex formation: a molecular double take? *Nat Struct Mol Biol,* 21**,** 20-5.
YASUDA, S. & HIROTA, Y. 1977. Cloning and Mapping of Replication Origin of Escherichia-Coli. *Proceedings of the National Academy of Sciences of the United States of America,* 74**,** 5458-5462.
YEELES, J. T., DEEGAN, T. D., JANSKA, A., EARLY, A. & DIFFLEY, J. F. 2015. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature,* 519**,** 431-5.
YOON, H. J. & CAMPBELL, J. L. 1991. The Cdc7 Protein of Saccharomyces-Cerevisiae Is a Phosphoprotein That Contains Protein-Kinase Activity. *Proceedings of the National Academy of Sciences of the United States of America,* 88**,** 3574-3578.
YOU, Z. & MASAI, H. 2008. Cdt1 forms a complex with the minichromosome maintenance protein (MCM) and activates its helicase activity. *J Biol Chem,* 283**,** 24469-77.
YU, J., FLEMING, S. L., WILLIAMS, B., WILLIAMS, E. V., LI, Z., SOMMA, P., RIEDER, C. L. & GOLDBERG, M. L. 2004. Greatwall kinase: a nuclear protein required

for proper chromosome condensation and mitotic progression in Drosophila. *J Cell Biol,* 164**,** 487-92.

ZACHARIAE, W., SCHWAB, M., NASMYTH, K. & SEUFERT, W. 1998. Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. *Science,* 282**,** 1721-1724.

ZEGERMAN, P. & DIFFLEY, J. F. X. 2007. Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature,* 445**,** 281-285.

ZENG, K., BASTOS, R. N., BARR, F. A. & GRUNEBERG, U. 2010. Protein phosphatase 6 regulates mitotic spindle formation by controlling the T-loop phosphorylation state of Aurora A bound to its activator TPX2. *J Cell Biol,* 191**,** 1315-32.

ZHAO, C., TOVAR, C., YIN, X., XU, Q., TODOROV, I. T., VASSILEV, L. T. & CHEN, L. 2009. Synthesis and evaluation of pyrido-thieno-pyrimidines as potent and selective Cdc7 kinase inhibitors. *Bioorg Med Chem Lett,* 19**,** 319-23.

ZHAO, F., ILBERT, M., VARADAN, R., CREMERS, C. M., HOYOS, B., ACIN-PEREZ, R., VINOGRADOV, V., COWBURN, D., JAKOB, U. & HAMMERLING, U. 2011. Are zinc-finger domains of protein kinase C dynamic structures that unfold by lipid or redox activation? *Antioxid Redox Signal,* 14**,** 757-66.

ZHONG, J., LIAO, J., LIU, X., WANG, P., LIU, J., HOU, W., ZHU, B., YAO, L., WANG, J., LI, J., STARK, J. M., XIE, Y. & XU, X. 2011. Protein phosphatase PP6 is required for homology-directed repair of DNA double-strand breaks. *Cell Cycle,* 10**,** 1411-9.